# آمار و احتمالات مهندسی

## جلسه یازدهم

# تحلیل دو متغیره

در این بخش به بررسی توزیع توامان دو متغیر تصادفی X و Y به طور همزمان با یکدیگر پرداخته می شود و این بررسی تنها زمانی اهمیت دارد که این دو متغیر تصادفی به نحوی با یکدیگر ارتباط داشته باشند و برای مواردی که X و Y از همدیگر مستقل هستند، تحلیل دومتغیره بی معناست.

در این بررسی ها فضای نمونه دو بعدی است، بنابراین مقادیر متغیر تصادفی در این بخش بایستی به صورت زوج مرتب باشند.

**مثال: سه بار یک سکهٔ سالم پرتاب می شود، X تعداد شیرها در سه پرتاب و Y تعداد شیرها در دو پرتاب اول.**

| فضای نمونه دوبعدی | TTT | HTT | TTH | THT | HHT | HTH | THH | HHH |
|---|---|---|---|---|---|---|---|---|
| (X,Y) | (0,0) | (1,1) | (1,0) | (1,1) | (2,2) | (2,1) | (2,1) | (3,2) |

# تحلیل دو متغیره

تابع احتمال توام (گسسته):

$$P(X=x, Y=y)=P[(X, Y)=(x,y)]$$

تابع چگالی توام (پیوسته):

$$f_{X,Y}(x,y)$$

تابع احتمال (یا چگالی توام Joint Density) برای **مثال قبل** عبارت است از:

| فضای نمونه دوبعدی | TTT | HTT | TTH | THT | HHT | HTH | THH | HHH |
|---|---|---|---|---|---|---|---|---|
| (X,Y) | (0,0) | (1,1) | (1,0) | (1,1) | (2,2) | (2,1) | (2,1) | (3,2) |

| Y \ X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 1/8 | 1/8 | 0 | 0 |
| 1 | 0 | 2/8 | 2/8 | 0 |
| 2 | 0 | 0 | 1/8 | 1/8 |

**مجموع این احتمالات نیز بایستی برابر با ۱ شود**

# تحلیل دو متغیره

## چگالی حاشیه ای (Marginal Density):

متغیر تصادفی پیوسته:

$$P(Y=y)=\sum_x P(X=x,Y=y) \qquad P(X=x)=\sum_y P(X=x,Y=y)$$

در مثال قبل چگالی حاشیه ای عبارت است از:

| Y \ X | 0 | 1 | 2 | 3 | $f_Y(y)$ |
|---|---|---|---|---|---|
| 0 | 1/8 | 1/8 | 0 | 0 | 2/8 |
| 1 | 0 | 2/8 | 2/8 | 0 | 4/8 |
| 2 | 0 | 0 | 1/8 | 1/8 | 2/8 |
| $f_X(x)$ | 1/8 | 3/8 | 3/8 | 1/8 | 1 |

متغیر تصادفی ناپیوسته:

$$f_X(x)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx$$

# تحلیل دو متغیره

## استقلال دو متغیر تصادفی:

در متغیرهای تصادفی X و Y، اگر احتمال توام متغیرها با حاصل ضرب چگالی حاشیه ای آنها برابر باشد، می توان گفت: دو متغیر مستقل از هم اند.

گسسته: $P(X=x, Y=y) = P(X=x)P(Y=y)$

پیوسته: $f_{X,Y}(x,y) = f_X(x).f_Y(y)$

**مثال: در دو بار پرتاب یک سکهٔ سالم، X تعداد شیرها و Y تعداد خط ها، جدول توزیع توام x و y بدین صورت است:**

| Y \ X | 0 | 1 | 2 | $f_Y(y)$ |
|---|---|---|---|---|
| 0 | 0 | 0 | ¼ | ¼ |
| 1 | 0 | 2/4 | 0 | 2/4 |
| 2 | ¼ | 0 | 0 | ¼ |
| $f_X(x)$ | ¼ | 2/4 | ¼ | 1 |

$$P(X=1, Y=1) = \frac{2}{4}$$

$$\left.\begin{array}{l} P(X=1) = \frac{1}{2} \\ P(Y=1) = \frac{1}{2} \end{array}\right\} \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

$$\boxed{\frac{2}{4} \neq \frac{1}{4}}$$

پس مستقل نیستند

# تحلیل دو متغیره

**کواریانس (Covariance):** به عنوان یک شاخص برای نشان دادن میزان ارتباط بین دو متغیر بکار می رود و رابطهٔ آن به صورت زیر است.

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)] = E(XY) - E(X)E(Y)$$

چنانچه دو متغیر تصادفی از یکدیگر مستقل باشند، داریم:

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0$$

یعنی در شرایط استقلال دو متغیر تصادفی **X** و **Y**، امید ریاضی حاصلضرب این دو متغیر **E(XY)** با حاصلضرب امید ریاضی آنها **E(X).E(Y)** برابر است.

اگر وابستگی بین دو متغیر وجود داشته باشد، مقدار کواریانس از صفر فاصله خواهد گرفت و مقادیر مثبت یا منفی به خود می گیرد.

$$Cov(X,X) = Var(X) \longrightarrow E(X.X) - E(X).E(X) = E(X^2) - E(X)^2 = Var(X)$$

$$Cov(X,Y) = Cov(Y,X)$$

$$Cov(X,c) = 0$$ مقدار ثابتی است $C$ $\longrightarrow$ $E(X.c) - E(X).E(c) = c.E(X) - E(X).c$

# تحلیل دو متغیره

## کواریانس (Covariance):

در صورت وجود رابطهٔ خطی بین دو متغیر $X$ و $Y$ به صورت $Y=aX+b$، کواریانس آنها به صورت زیر است:

$$Cov(X,Y)=aCov(X,X)=aVar(X)$$

$$E(X.(aX+b))-E(X).E(aX+b)=E(aX^2+bX)-E(X).(aE(X)+b)$$
$$=aE(X^2)+bE(X)-aE(X)^2-bE(X)=a[E(X^2)-E(X)^2]=a.Var(X)$$

**مثال:** سه بار یک سکهٔ سالم پرتاب می شود، $X$تعداد شیرها در سه پرتاب و $Y$تعداد شیرها در دو پرتاب اول.

| فضای نمونه دوبعدی | TTT | HTT | TTH | THT | HHT | HTH | THH | HHH |
|---|---|---|---|---|---|---|---|---|
| (X,Y) | (0,0) | (1,1) | (1,0) | (1,1) | (2,2) | (2,1) | (2,1) | (3,2) |

| Y \ X | 0 | 1 | 2 | 3 | $f_Y(y)$ |
|---|---|---|---|---|---|
| 0 | 1/8 | 1/8 | 0 | 0 | 2/8 |
| 1 | 0 | 2/8 | 2/8 | 0 | 4/8 |
| 2 | 0 | 0 | 1/8 | 1/8 | 2/8 |
| $f_X(x)$ | 1/8 | 3/8 | 3/8 | 1/8 | 1 |

# تحلیل دو متغیره

## کواریانس (Covariance):

**مثال:** سه بار یک سکهٔ سالم پرتاب می شود، X تعداد شیرها در سه پرتاب و Y تعداد شیرها در دو پرتاب اول.

| فضای نمونه دوبعدی | TTT | HTT | TTH | THT | HHT | HTH | THH | HHH |
|---|---|---|---|---|---|---|---|---|
| (X,Y) | (0,0) | (1,1) | (1,0) | (1,1) | (2,2) | (2,1) | (2,1) | (3,2) |

| Y \ X | 0 | 1 | 2 | 3 | $f_Y(y)$ |
|---|---|---|---|---|---|
| 0 | 1/8 | 1/8 | 0 | 0 | 2/8 |
| 1 | 0 | 2/8 | 2/8 | 0 | 4/8 |
| 2 | 0 | 0 | 1/8 | 1/8 | 2/8 |
| $f_X(x)$ | 1/8 | 3/8 | 3/8 | 1/8 | 1 |

کواریانس $X$ و $Y$ را محاسبه کنید:

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$E(XY) = \sum_X \sum_Y y.x.P(X=x, Y=y) = 0\times 0\times \frac{1}{8} + 1\times 0\times 0 + 2\times 0\times 0 + 0\times 1\times \frac{1}{8} + \cdots$$

$$+1\times 1\times \frac{2}{8} + 2\times 1\times 0 + 0\times 2\times 0 + 1\times 2\times \frac{2}{8} + 2\times 2\times \frac{1}{8} + 0\times 3\times 0 + 1\times 3\times 0 + 2\times 3\times \frac{1}{8}$$

$$= {}^{2}/_{8} + {}^{4}/_{8} + {}^{4}/_{8} + {}^{6}/_{8} = 2$$

$$E(X) = 0\times \frac{1}{8} + 1\times \frac{3}{8} + 2\times \frac{3}{8} + 3\times \frac{1}{8} = \frac{12}{8}$$

$$E(Y) = 0\times \frac{2}{8} + 1\times \frac{4}{8} + 2\times \frac{2}{8} = 1$$

$$E(X)\times E(Y) = \frac{12}{8}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 2 - \frac{12}{8}\times 1 = \frac{4}{8} = 0.5$$

# تحلیل دو متغیره

## کواریانس (Covariance):

**مسئله:**

اگر $X$ و $Y$ دو متغیر تصادفی باشند، **نشان دهید** کواریانس $cY+d$ و $aX+b$ برابر است با $a.c$ برابر کواریانس $X$ و $Y$.

$$Cov(aX+b, cY+d) = a.c.Cov(X,Y)$$

# تحلیل دو متغیره

## ضریب همبستگی (Correlation Coefficient):

چنانچه واحد اندازه گیری متغیرهای تصادفی توامان با یکدیگر متفاوت باشد، مقادیر بدست آمده برای امید ریاضی توامان و کواریانس آنها اعداد معنی داری نیست، زیرا واحدهای اندازه گیری متفاوتی با یکدیگر دارند. استفاده از ضریب همبستگی برای چنین مواردی مناسب است. در واقع ضریب همبستگی میزان ارتباط خطی بین دو متغیر تصادفی توامان را بدون در نظر گرفتن واحدهای آنها بیان می کند. از اینرو خود ضریب همبستگی فاقد واحد است.

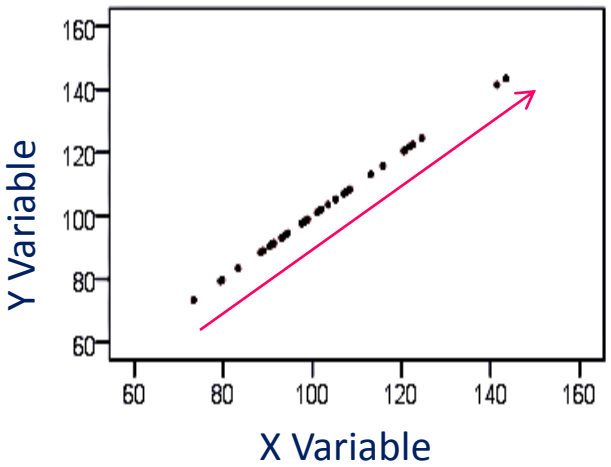
در اندازه گیری کواریانس، چنانچه مقدار کواریانس مثبت بدست می آمد بیانگر وجود رابطهٔ مستقیم بین دو متغیر بود و زمانی که منفی بدست می آمد بیانگر وجود رابطهٔ معکوس بود. ضریب همبستگی نیز چنین عمل می کند.

# تحلیل دو متغیره

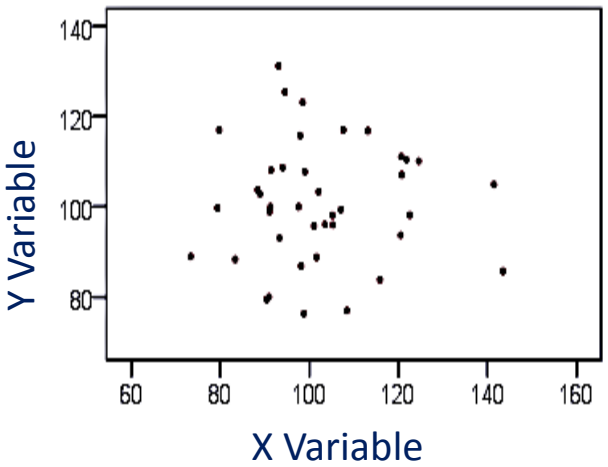
## ضریب همبستگی (Correlation Coefficient):

مقدار این ضریب در بازهٔ (1+ ,1-) تغییر می کند.

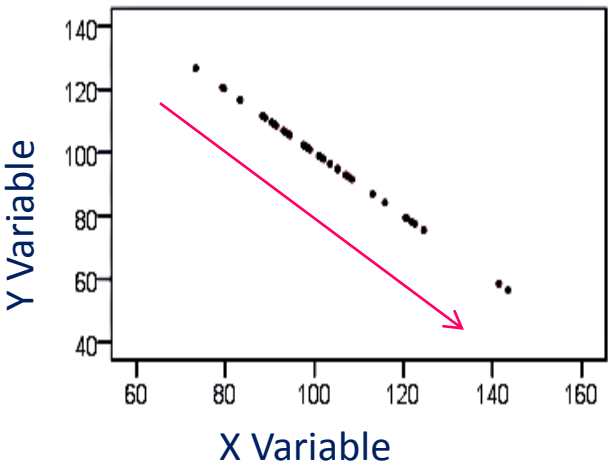
ضریب همبستگی مثبت (۱+)




ضریب همبستگی صفر (۰)




ضریب همبستگی منفی(۱-)




$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

| وجود رابطه خطی مستقیم | عدم وجود رابطه خطی | وجود رابطه خطی معکوس |
|---|---|---|
| +1 | 0 | -1 |

# • انواع همبستگی

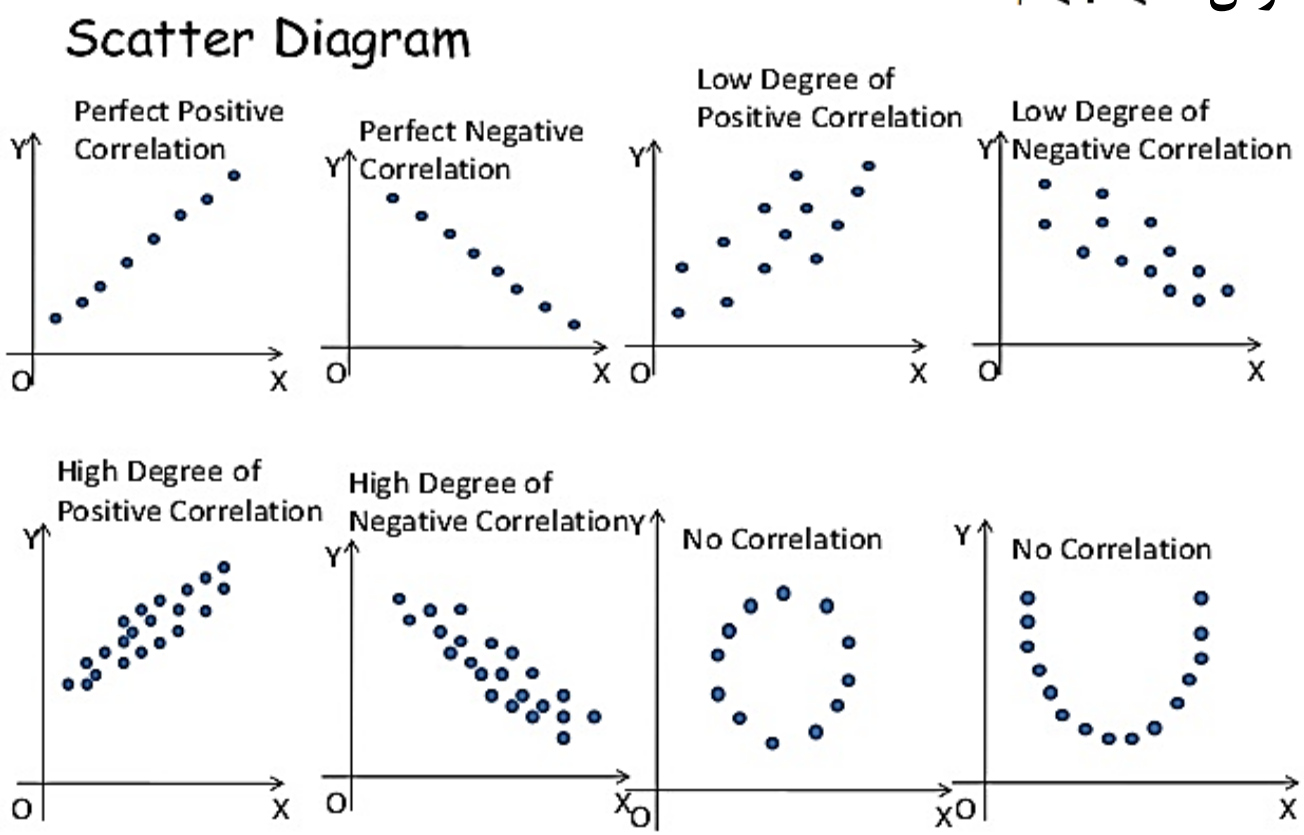
۱- همبستگی کامل و مثبت (مستقیم) (شیب مثبت) r=۱

۲- همبستگی ناقص و مثبت (مستقیم) ۰<r<۱

۳- همبستگی کامل و معکوس (شیب منفی) r=-۱

۴- همبستگی ناقص و معکوس ۰>r>-۱

۵- عدم همبستگی r=0



# تحلیل دو متغیره

## ضریب همبستگی (Correlation Coefficient):

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

ضریب همبستگی برای مثال پرتاب سه سکه برابر است با:

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 2 - \frac{12}{8} = \frac{4}{8} = \frac{1}{2}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{13}{4} - \frac{144}{64} = 1$$

$$Var(Y) = E(Y^2) - E(Y)^2 = \frac{12}{8} - 1 = \frac{1}{2}$$

$$Corr(X,Y) = \frac{1/2}{\sqrt{1/2}} = \frac{\sqrt{2}}{2} \approx 0.71$$

ضریب همبستگی با تغییر **مقیاس** (یا ضرب یک مقدار ثابت در مقادیر متغیر تصادفی) و تغییر **مکان** (یا جمع یک مقدار ثابت با مقادیر متغیر تصادفی) تغییر نخواهد کرد.

$$Corr(aX+b, cY+d) = Corr(X, Y)$$

# تحلیل دو متغیره

- فرمول های محاسبه ضریب همبستگی پیرسون :

(1) $$r_{X,Y} = \operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y},$$

(2) $$r_{X,Y} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^۲ - (\sum X)^۲][n\sum Y^۲ - (\sum Y)^۲]}}$$

(3) $$r_{X,Y} = \frac{\sum XY - n\bar{X}\,\bar{Y}}{\sqrt{[\sum X^۲ - n\bar{X}^۲][\sum Y^۲ - n\bar{Y}^۲]}}$$

(4) $$r_{X,Y} = \frac{\sum (X-\bar{X})(Y-\bar{Y})}{\sqrt{(\sum (X-\bar{X})^۲ \sum (Y-\bar{Y})^۲}}$$

با توجه به این که صورت سوال چه اطلاعاتی را در اختیار می گذارد، یکی از روابط بالا استفاده می شود.

# تحلیل دو متغیره

## ضریب همبستگی (Correlation Coefficient):

مثال) از یک جامعه دو بعدی، پس از محاسبات لازم کمیت های زیر به دست آمده است:

$n=۵, \quad \sum X_i = ۱۲۵, \quad \sum X_i^۲ = ۶۵۰, \quad \sum Y_i = ۱۰۰, \quad \sum Y_i^۲ = ۴۳۶, \quad \sum X_i Y_i = ۵۴۰$

ضریب همبستگی خطی بین دو صفت X و Y چقدر است؟

با توجه به اطلاعات داده شده در صورت سؤال بهترین فرمول دومی است:

$$r_{X,Y} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^۲ - (\sum X)^۲][n\sum Y^۲ - (\sum Y)^۲]}} = \frac{(۵)(۵۴۰)-(۱۲۵)(۱۰۰)}{\sqrt{[(۵)(۶۵۰)-(۱۲۵)^۲][(۵)(۴۳۶)-(۱۰۰)^۲]}} = \frac{-۹۸۰۰}{۹۸۳۷/۳}$$

$$r_{X,Y} = -۰/۹۹$$

# ویژگی های ضریب همبستگی

۱) $-۱ \leq r \leq ۱$

۲) ضریب همبستگی ماهیتاً قرینه است: $r_{xy} = r_{yx}$

۳) ضریب همبستگی مستقل ازمبدأ و مقیاس اندازه گیری است.

۴) اگر x و y مستقل باشند، ضریب همبستگی برابر صفراست اما اگر ضریب همبستگی برابر صفر باشد الزاماً دو متغیرمستقل نیستند.

۵) ضریب همبستگی معیاری است جهت همبستگی خطی یا وابستگی خطی و برای توصیف ارتباطات غیرخطی به کار نمی رود.

۶) ضریب همبستگی الزاماً بیانگر هیچ گونه رابطه علت و معلولی نیست.

## • نکاتی در مورد ضریب همبستگی

-وقتی از ضریب همبستگی صحبت می کنیم، نمی توان گفت که مثبت بهتر است یا منفی. به طور کلی هر چه قدر عدد به دست آمده صرفنظر از علامت آن، بزرگتر باشد، همبستگی بیشتر است و علامت مثبت و منفی فقط بر جهت همبستگی دلالت دارد.

-اگر r= ۰/۳ یا r= ۰/۹ باشد، این فقط بدان معنا است که ضریب همبستگی در دو حالت مثبت است و نمی توان گفت که r= ۰/۹ سه برابر همبستگی نسبت به r= ۰/۳ است.

-اگر همبستگی بین دو متغیر صفر باشد، این بدین معنی است که این دو متغیر با هم رابطه خطی ندارند. اما ممکن است به طور غیر خطی به یکدیگر مربوط باشند.

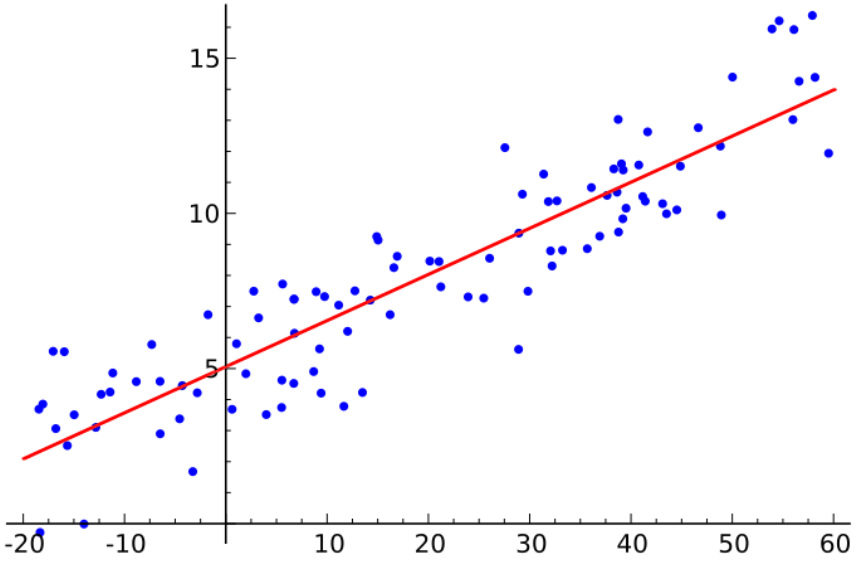
- اگر تمام اعداد یا نمره ها با هم برابر باشند یعنی پراکندگی صفر است و ضریب همبستگی نیز صفر است.

# تحلیل دو متغیره

## خط برگشت (Regression):

در تحلیل همبستگی هدف اولیه اندازه گیری میزان همبستگی خطی بین دو متغیر است اما اصولاً در تحلیل رگرسیون به دنبال چنین اندازه گیری نیستیم بلکه سعی داریم مقدار متوسط یک متغیر را بر اساس مقادیر ثابت متغیرهای دیگر تخمین بزنیم یا پیش بینی کنیم.

ما باید فقط وقتی از معادله خط رگرسیون استفاده کنیم که r یک همبستگی خطی معنادار را نشان می دهد. به طور کلی اگر r نزدیک به ۱ یا ۱- باشد، آنگاه خط رگرسیون بر داده ها برازنده است اما اگر r نزدیک صفر باشد، برازندگی خط رگرسیون کم خواهد بود.



# تحلیل دو متغیره

## خط برگشت (Regression):

رگرسیون ارتباط تابعی بین دو متغیر را نشان می دهد. در رگرسیون یک متغیر مستقل $X$ وجود دارد که البته تصادفی نبوده و کاملاً قابل کنترل بوده و مقادیر آن از طریق محاسبات یا اندازه گیری حاصل می شود.

متغیر دیگری به نام $Y$ نیز وجود دارد که وابسته به مقدار متغیر $X$ است و به نوعی یک متغیر تصادفی است.

به عنوان مثال چنانچه نمرهٔ درس ریاضی دانش آموزان را متغیر مستقل $X$ در نظر بگیریم، درصد جوابهای درست درس ریاضی در کنکور را می توان $Y$ فرض کرد که اولاً به $X$ وابسته بوده و دوماً متغیر تصادفی است.

در نتیجه خط برگشت یا رگرسیون نشان دهندهٔ مدل ارتباطی این دو متغیر است.

# تحلیل دو متغیره

## خط برگشت (Regression):

همانطور که از نامش پیداست، هرگاه مدل ارتباطی بین دو متغیر مستقل و وابسته به صورت خطی باشد از رگرسیون استفاده می شود. فرمول کلی رابطهٔ بین $X$ و $Y$ به فرم روبرو است:

$$Y = \underbrace{aX + b}_{\text{غیر تصادفی}} + \underbrace{Error}_{\text{تصادفی}}$$

بنابراین خطای تصادفی (یا *Random Error*) برابر است با:

$$Error = Y - (aX + b)$$

که این نوع خطای تصادفی دارای میانگین صفر و واریانس $\sigma^2$ است.

بدین ترتیب پارامترهای مدل عبارتند از $a$، $b$ و $\sigma^2$ زیرا متغیر تصادفی $Y$ دارای میانگین (یا امید ریاضی) $aX+b$ و واریانس $\sigma^2$ است.

$$E(Y) = E(aX + b + Error) = aX + b + E(Error) = aX + b$$

$$Var(Y) = Var(aX + b + Error) = 0 + Var(Error) = \sigma^2$$

# تحلیل دو متغیره

## برآورد پارامترهای خط برگشت:

فرض می کنیم زوج های مرتب از یک نمونهٔ تصادفی $n$تایی به صورت زیر است:

$$(X_1\ Y_1), (X_2\ Y_2), (X_3\ Y_3) \ldots (X_n\ Y_n)$$

پس داریم:

$$Y_i = aX_i + b + E_i$$

که در آن $E_i$ها مستقل و دارای توزیع نرمال هستند.

$E(Y_i) = aX_i + b \qquad Var(Y_i) = \sigma^2$

روشی که در اینجا برای برآورد کردن پارامترهای مدل مورد استفاده قرار می گیرد بر اساس کمینه کردن مربع خطا (*Least Square Error*) است.

مجموع مربع خطاها (*Sum of Square Error*) را به صورت زیر بدست می آوریم:

$$SSE = \sum_{i=1}^{n} [Y_i - (aX_i + b)]^2 = A(a, b)$$

با کمینه کردن مقدار $SSE$ مقادیر پارامترهای $a$ و $b$ و $\sigma^2$ بدست می آیند.

# تحلیل دو متغیره

## برآورد پارامترهای خط برگشت:

بدیهی است ضرایب $a$ و $b$ بایستی طوری برآورد گردند که مجموع خطاهای تمام نقاط نمونه کمترین مقدار خود را داشته باشد. بنابراین از طریق مشتق گیری از تابع $SSE$ می توان به روابط زیر رسید:

$$a = \frac{S_{xy}}{S_x} \qquad b = \bar{y} - a\bar{x} \qquad \sigma^2 = (S_y)^2 - a^2(S_x)^2$$

$$S_x = \sum (x - \bar{x})^2 = \sum x^2 - \frac{1}{n}(\sum x)^2$$

$$S_y = \sum (y - \bar{y})^2 = \sum y^2 - \frac{1}{n}(\sum y)^2$$

$$S_{xy} = \sum (x - \bar{x})(y - \bar{y}) = \sum xy - \frac{1}{n}(\sum x)(\sum y)$$

# تحلیل دو متغیره/ رگرسیون

## مثال:

فرض کنید مقادیر $X$ و $Y$ به صورت جدول زیر است. رابطهٔ خطی بین $X$ و $Y$ را بدست آورید.

| $X$ | $Y$ |
|---|---|
| 35 | 9 |
| 49 | 15 |
| 21 | 7 |
| 39 | 11 |
| 15 | 5 |
| 28 | 8 |
| 25 | 9 |

$$n = 7$$

$$\bar{x} = \frac{\sum x}{n} = \frac{212}{7} = 30.2857$$

$$\bar{y} = \frac{\sum y}{n} = \frac{64}{7} = 9.1429$$

$$S_x = \sum x^2 - \frac{1}{n}(\sum x)^2 = 7222 - \frac{(212)^2}{7} = 801.4286$$

$$S_{xy} = \sum xy - \frac{1}{n}(\sum x)(\sum y) = 2150 - \frac{(212)(64)}{7} = 211.7143$$

$$a = \frac{S_{xy}}{S_x} = 0.2642 \qquad b = \bar{y} - a\bar{x} = 1.1414$$

$$y = 0.2642x + 1.1414$$

# ضریب تعیین یا ضریب تشخیص ($r^2$) (Coefficient of Determination)

- $r^2$ نسبت یا درصدی از تغییر کل در متغیر وابسته y را که توسط متغیر x توضیح داده می شود، به دست می دهد.
- $r^2$ نشان می دهد چند درصد از کل واریانس y ناشی از واریانس x است.
- $r^2$ برآوردی از واریانس مشترک بین دو متغیر را نشان می دهد.
- $r^2$ میزان اشتراک تغییرات است.
- $r^2$ درصد یا نسبتی است که تغییرات متغیر وابسته y میتواند توسط تغییرات در متغیر x توضیح داده شود و $1-r^2$ درصد یا نسبتی است که تغییرات y توسط تغییرات x توضیح داده نمی شود.
- $r^2$ کمیتی غیر منفی است و حدودش قطعاً کوچکتر مساوی ۱ و بزرگتر مساوی صفر است. یعنی: $0 \le r^2 \le 1$

در مسائل رگرسیون،ضریب همبستگی با جذر گرفتن از ضریب تعیین به دست می آید:

$$r = \pm \sqrt{r^2}$$

# ضریب تعیین یا ضریب تشخیص ($r^۲$) (Coefficient of Determination)

در مسائل رگرسیون،ضریب همبستگی با جذر گرفتن از ضریب تعیین به دست می آید:

$$r = \pm\sqrt{r^۲}$$

نکته مهم: ضریب تعیین برای مقایسه ضریب های همبستگی مفید است. وقتی فردی یک r به مقدار ۰/۸ را با r دیگری به مقدار ۰/۴ مقایسه می کند، این تمایل وجود دارد که ۰/۸ را دو برابر ۰/۴ بداندانند اما این مقایسه صحیح نیست. ضریب های همبستگی را باید بر حسب تغییر مشترک دو متغیر (ضریب تشخیص) مقایسه کرد:

$r_۱ = ۰/۸$ $\qquad r_۱^۲ = ۰/۶۴$

$r_۲ = ۰/۴$ $\qquad r_۲^۲ = ۰/۱۶$

$\rightarrow \frac{۰/۶۴}{۰/۱۶} = ۴$ $r_۱$ چهار برابر $r_۲$ است.