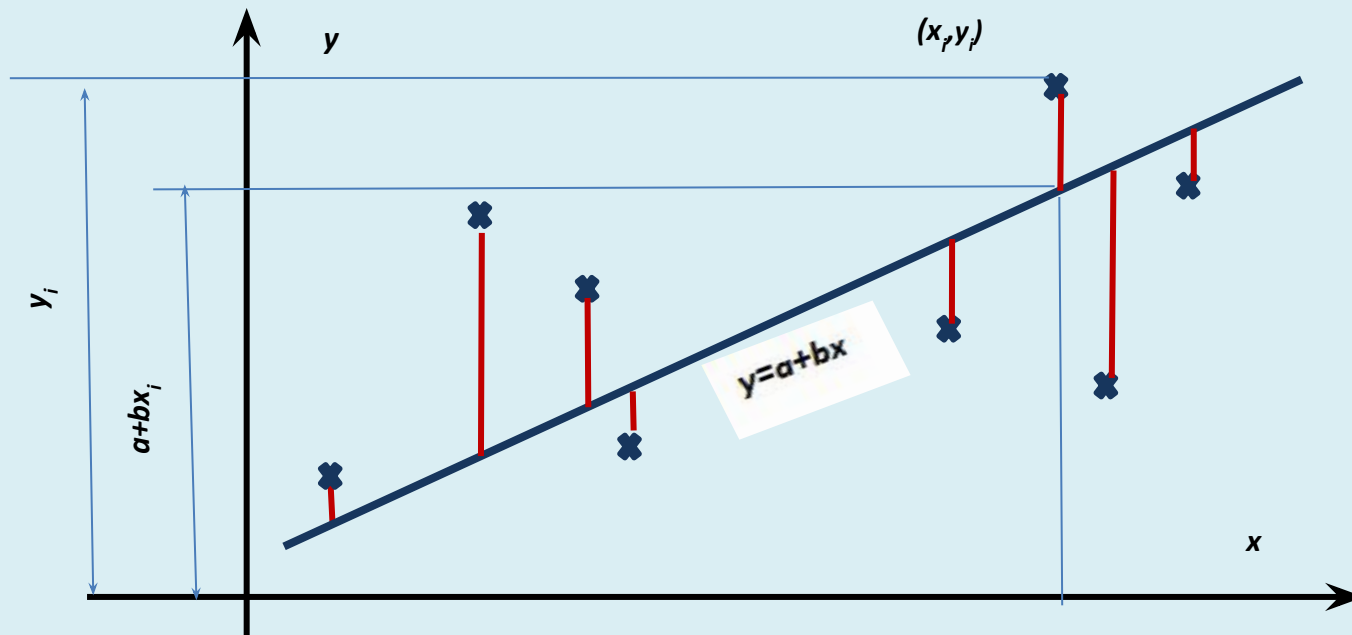


# Сызықтық жұп регрессия теңдеуі

Байқаудан (сынақтан, тәжірибеден) мынандай мәлімет алдық делік:

X тәуелсіз айнымалы (фактор)	$x_1$	$x_2$	...	$x_n$
$Y = \tilde{y}(x) + \varepsilon$ тәуелді айнымалы, қорытынды, нәтиже	$y_1$	$y_2$	...	$y_n$

Регрессия теңдеуін  $y = a + bx$  түрінде (яғни регрессия сызығын түзу деп есептейік) қарастыралық. Енді таблицандағы шамаларды пайдаланып  $a, b$  параметрлерінің баламасын табуымыз керек, яғни басқаша айтқанда  $a, b$  шамаларын табуымыз керек.



Осы параметрлерді табу үшін мынадай шаманы қарастыралық:

$$S = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

Бұл қосынды кіші квадраттар әдісінде қолданылатын қосынды.

Түзу берілген нүктелерге мейлінше жақын орналасуы үшін осы қосынды өзінің ең кіші мәнін қабылдауы керек, яғни минимумын. Сондықтан оның  $a$ ,  $b$  бойынша туындыларын тауып нольге теңестіріп, әртүрлі түрлендірулерден кейін мына алгебралық теңдеулер системасын аламыз:

$$\begin{cases} an + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i; \\ a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i. \end{cases}$$

Осы теңдеулер системасының шешімі мынандай болады:

$$b = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}, \quad a = \bar{y} - b\bar{x}.$$

$$b = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}, \quad a = \bar{y} - b\bar{x}.$$

**Формулаларда мынандай белгілеулер қолданылған:**

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n}, \quad \overline{x \cdot y} = \frac{\sum_{i=1}^n x_i y_i}{n}, \quad \overline{x^2} = \frac{\sum_{i=1}^n x_i^2}{n}.$$

$$\text{cov}(x, y) = \overline{x \cdot y} - \bar{x} \cdot \bar{y}, \quad \sigma_x^2 = \overline{x^2} - \bar{x}^2.$$

Регрессия теңдеуіндегі  $b$  параметрін регрессия коэффициенті деп атайды. Аналитикалық геометрияда біз оны түзудің бұрыштық коэффициенті деп атаған болатынбыз. Бұл коэффициент  $x$  факторы бір бірлікке өзгергенде  $y$  қанша бірлікке өзгертетінін көрсетеді.

Excel электрондық таблицасында  $b$  параметрін табу үшін НАКЛОН функциясын қолдануға болады.

Регрессия теңдеуінің екінші параметрі  $a$  регрессия түзуінің  $Oy$  осін қай нүктеде кесіп өтетінін көрсетеді. Excel электрондық таблицасында бұл параметрді табу үшін ОТРЕЗОК функциясын қолдануға болады.

Жалпы Excel электрондық таблицасын қолданып регрессия теңдеуін бірнеше әдіспен табуға болады:

- Корреляция өрісінің графигін салып болғасын оның тенденциясын салған кезде теңдеуді көрсету керек деген нұсқау берілсе, графикте теңдеу көрсетіледі.
- “Данные -Анализ данных –Регрессия” командасын қолдану арқылы регрессия параметрлерін табуға болады. Бұл туралы лабораториялық сабақта толық айтамыз.

## Регрессия теңдеуін табу мысалы

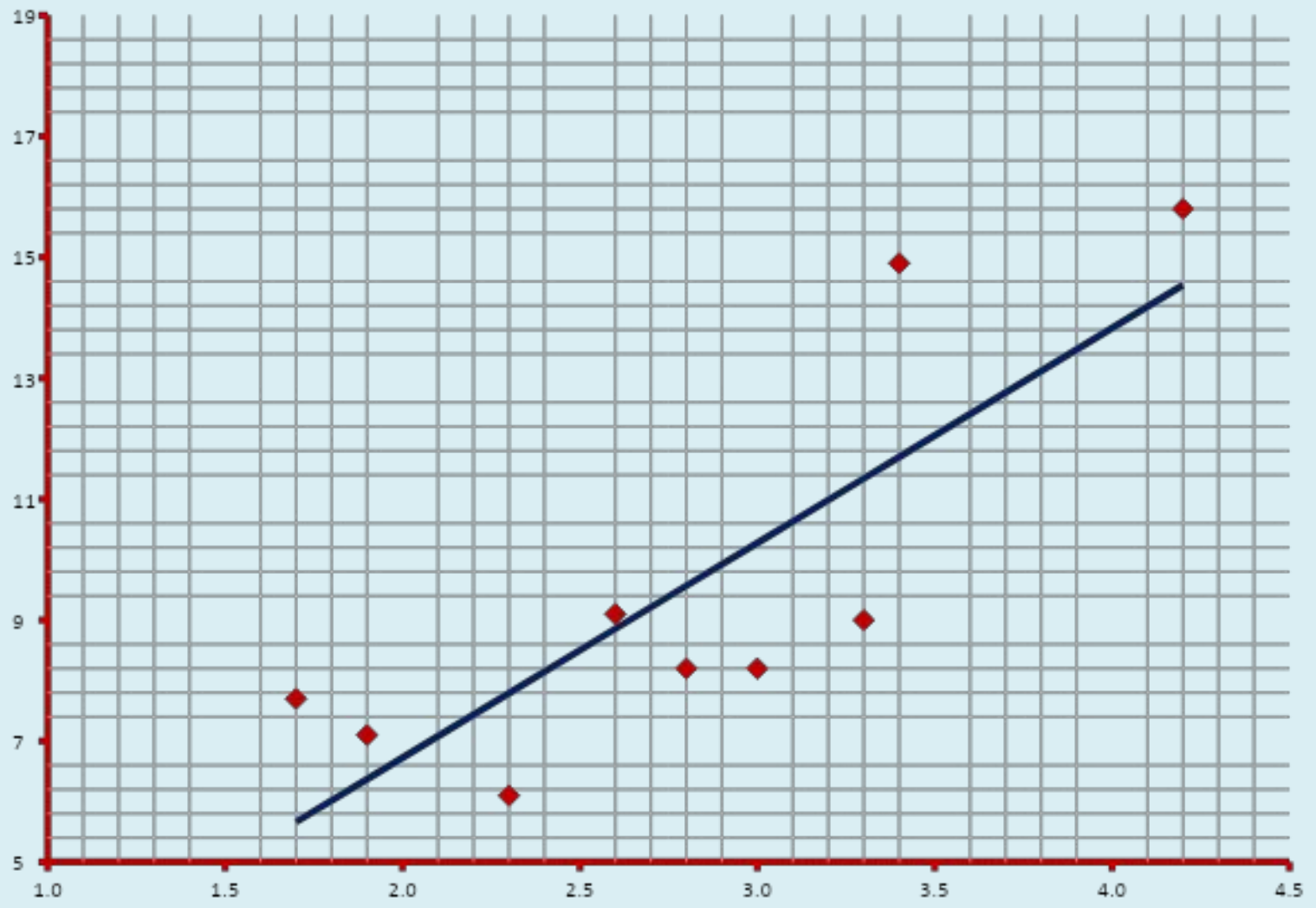
Төмендегі таблицада 9 сауда фирмасындағы бір айда сатылған мата көлемі  $x$  (мың м) пен айлық сый ақы  $y$  (мың сом) көрсетілген.

	1	2	3	4	5	6	7	8	9
$x$	2.3	2.8	1.9	3.4	2.6	3.3	4.2	3.0	1.7
$y$	6.1	8.2	7.1	14.9	9.1	9.0	15.8	8.2	7.7

Сатылған мата мен сый ақы көлемдерінің арасындағы байланысты зерттеу керек. Регрессия теңдеуін табыңыз.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$b = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}, \quad a = \bar{y} - b\bar{x}.$$



*A. All cases*

<b>N</b>	<b>x</b>	<b>y</b>	<b>xy</b>	<b>x<sup>2</sup></b>	$x - \bar{x}$	$(x - \bar{x})^2$	$y - \bar{y}$	$(y - \bar{y})^2$
<b>1</b>	<b>2.3</b>	<b>6.1</b>	<b>14.0</b>	<b>5.3</b>	<b>-0.5</b>	<b>0.25</b>	<b>-3.47</b>	<b>12.02</b>
<b>2</b>	<b>2.8</b>	<b>8.2</b>	<b>23.0</b>	<b>7.8</b>	<b>0.0</b>	<b>0.00</b>	<b>-1.37</b>	<b>1.87</b>
<b>3</b>	<b>1.9</b>	<b>7.1</b>	<b>13.5</b>	<b>3.6</b>	<b>-0.9</b>	<b>0.81</b>	<b>-2.47</b>	<b>6.08</b>
<b>4</b>	<b>3.4</b>	<b>14.9</b>	<b>50.7</b>	<b>11.6</b>	<b>0.6</b>	<b>0.36</b>	<b>5.33</b>	<b>28.44</b>
<b>5</b>	<b>2.6</b>	<b>9.1</b>	<b>23.7</b>	<b>6.8</b>	<b>-0.2</b>	<b>0.04</b>	<b>-0.47</b>	<b>0.22</b>
<b>6</b>	<b>3.3</b>	<b>9.0</b>	<b>29.7</b>	<b>10.9</b>	<b>0.5</b>	<b>0.25</b>	<b>-0.57</b>	<b>0.32</b>
<b>7</b>	<b>4.2</b>	<b>15.8</b>	<b>66.4</b>	<b>17.6</b>	<b>1.4</b>	<b>1.96</b>	<b>6.23</b>	<b>38.85</b>
<b>8</b>	<b>3.0</b>	<b>8.2</b>	<b>24.6</b>	<b>9.0</b>	<b>0.2</b>	<b>0.04</b>	<b>-1.37</b>	<b>1.87</b>
<b>9</b>	<b>1.7</b>	<b>7.7</b>	<b>13.1</b>	<b>2.9</b>	<b>-1.1</b>	<b>1.21</b>	<b>-1.87</b>	<b>3.48</b>
<b>Қосынды</b>	<b>25.2</b>	<b>86.1</b>	<b>258.6</b>	<b>75.5</b>	<b>0.0</b>	<b>4.9</b>	<b>0.0</b>	<b>93.2</b>
<b>Орта мән</b>	<b>2.8</b>	<b>9.6</b>	<b>28.7</b>	<b>8.4</b>	<b>0.0</b>	<b>0.5</b>	<b>0.0</b>	<b>10.4</b>

**Формулаға қойып корреляция коэффициентін табамыз**

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{17,47}{\sqrt{4,92 \times 93,16}} = 0,816$$

**Сызықтық корреляция коэффициенті 1 санына жақын, сондықтан айнымалылар арасында тығыз тәуелділік бар деп айтуға болады (егер корреляция коэффициентінің абсолют шамасы 0,7 санынан үлкен болса тәуелділік тығыз деп есептеуге болады)**

**Енді регрессия теңдеуінің параметрлерін есептелік**

$$b = \frac{\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y}}{\bar{x}^2 - \bar{x}^2} = \frac{28,7 - 2,8 \times 9,6}{8,4 - 2,8 \times 2,8} = 3,551.$$

$$a = \bar{y} - b\bar{x} = 9,6 - 3,551 \times 2,8 = -0,376$$

**Сонда регрессия теңдеуі мынадай болады екен**

$$y = -0,376 + 3,551x$$



**Біз әзірге тек қана регрессияның сызықтық теңдеуін қалай табуды ғана көрсеттік. Бірақ регрессия теңдеуін экономикада болжам жасау үшін біз әуелі осы тапқан теңдеуімізге қаншалықты сенуге болатынын көрсетуәміз керек, егер ол сенімділік шамасы бізді қанағаттандыратын болса, онда оны экономикада болжауға қолданамыз.**

**Регрессия теңдеуіне сенуге болатынын білу үшін оның параметрлерінің және теңдеудің өзінің маңыздылығын білуіміз керек. Ол туралы келесі сабақтарда айтылады.**

**Сонымен қатар, болжау қолданған кезде ол болжаудан алынған мәннің сенімділік аралығын білу керек болады. Оны есептеудің де арнайы формуласы бар.**

**Егер сынақ саны тым көп болса, онда тәуелсіз айнымалы мен тәуелді айнымалының сынақтан алынған мәндері бірнеше рет қайталануы мүмкін, ондай жағдайда біз көрсеткен есептеу әдісімен есептеу өте көп ұзаққа созылып кетеді. Бұл жағдайда басқаша есептеу алгоритмі қолданылады. Ол туралы кейін айтылады.**