Здравствуйте!

Числовые характеристики случайных величин

Выделяют 3 основные группы: **характеристики положения, характеристики разброса и характеристики связи.**

Характеристики положения определяют некоторое среднее числовое значение случайной величины, в окрестности которого наиболее часто встречается большинство ее значений.

Характеристики разброса определяют величину, меру разброса, отклонения значений случайной величины от ее среднего значения.

Характеристики связи определяют силу зависимости случайных величин, т.е. как значение некоторой случайной величины влияет на значения других случайных величин.

Математическое ожидание и медиана

Математическое ожидание (среднее значение) и медиана относятся к так называемым **характеристикам положения** случайной величины, цель которых – указать **число,** около которого группируются остальные значения случайной величины. Это число является как бы наиболее типичным значением случайной величины.

Пусть заданы: вероятностное пространство $\{\Omega, F, P\}$, случайная величина $\xi(\omega)$ со своей функцией распределения $F_\xi(x)$ и борелевская функция $f(x)$, тогда математическим ожидание величины $\eta = f(\xi)$ называется число $M\{f(\xi)\}$, определяемое равенством $M\{f(\xi)\} = \int\limits_{\Omega} f(\xi(\omega)) dP(\omega) = \int\limits_{-\infty}^{\infty} f(x) dF_\xi(x)$

В частности, математическим ожиданием случайной величины $\xi(\omega)$ называется число $M\{\xi(\omega)\} = \int\limits_{\Omega} \xi(\omega) dP(\omega) = \int\limits_{-\infty}^{\infty} x dF_\xi(x)$.

Вычисление математического ожидания определяется способами вычисления интеграла Стилтьеса.

Для <u>непрерывных</u> случайных величин, когда существует плотность распределения $p(x)=F'(x)$ непрерывной случайной величины ξ, получим

$$M\{f(\xi)\}=\int_{-\infty}^{\infty} f(x)\,dF_{\xi}(x)=\int_{-\infty}^{\infty} f(x)\,F'(x)dx=\int_{-\infty}^{\infty} f(x)\,p(x)dx,$$

$$M\{\xi\}=\int_{-\infty}^{\infty} x\,p(x)dx.$$

Для <u>дискретных</u> случайных величин, когда задан ряд распределения $\{x_i, p_i\}$

Значения случайной величины	x_1	x_2	…	x_k
вероятности	p_1	p_2	…	p_k

дискретной случайной величины ξ, запишем

$$M\{f(\xi)\}=\int_{-\infty}^{\infty} f(x)\,dF_{\xi}(x)=\sum_i f(x_i)\,p_i,$$

Определенное таким способом среднее значение случайной величины называют ***математическим ожиданием*** этой случайной величины: $M\{\xi\}=\sum_i x_i p_i$

Для случайного вектора $\xi = \{\xi_1, \xi_2, ..., \xi_n\}$, заданного функцией распределения $F(x) = F(x_1, x_2, ..., x_n)$ и борелевской функции $f(x) = f(x_1, x_2, ..., x_n)$ для математического ожидания запишем интеграл Лебега

$$M\{f(\xi)\} = M\{f(\xi_1, \xi_2, ..., \xi_n)\} = \int_{\Omega} f(\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega))\, dP(\omega),$$

который равен многомерному интегралу Стилтьеса

$$\int_{\Omega} f(\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega))\, dP(\omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} f(x_1, x_2, ..., x_n) F(dx_1, dx_2, ..., dx_n),$$

а этот многомерный интеграл Стилтьеса для непрерывных случайных величин при заданной плотности распределения $p(x_1, x_2, ..., x_n) = \dfrac{\partial^n F(x_1, x_2, ..., x_n)}{\partial x_1 \partial x_2 ... \partial x_n}$,

так же как одномерный, определяется интегралом Римана

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} f(x_1, x_2, ..., x_n) F(dx_1, dx_2, ..., dx_n) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} f(x_1, x_2, ..., x_n)\, p(x_1, x_2, ..., x_n)\, dx_1, dx_2,.$$

Для многомерных случайных величин многомерный интеграл Стилтьеса определяется многомерной суммой.

Итак, пусть дана случайная величина ξ. Тогда её математическое ожидание M{ξ} определяется следующим образом:

$$
\mathrm{M}\{\xi\} = \begin{cases} \sum\limits_{i=1}^{\infty} x_i p_i & \text{для дискретной случайной величины,} \\ \int\limits_{-\infty}^{\infty} x p_{\xi}(x) dx & \text{для непрерывной случайной величины.} \end{cases}
$$

В самом общем виде математическое ожидание записывается через интеграл Стилтьеса:

$$
\mathrm{M}\{\xi\} = \int\limits_{-\infty}^{\infty} x dF_{\xi}(x).
$$

Свойства математического ожидания

1. $\mathrm{M}\{c\} = c$.

Действительно, константу можно рассматривать как случайную величину, которая принимает значение c с вероятностью 1 и любое другое значение с вероятностью 0. Тогда

$$\mathrm{M}\{c\} = c \cdot 1 + (\text{любое другое значение}) \cdot 0 = c \text{ или}$$

$$M\{c\} = \int_{\Omega} c\, dP(\omega) = c \int_{\Omega} dP(\omega) = c P(\Omega) = c$$

2. $\mathrm{M}\{c\xi\} = c \cdot \mathrm{M}\{\xi\}$.

Действительно,

$$\mathrm{M}\{c\xi\} = \int_{-\infty}^{\infty} c x p_{\xi}(x) dx = c \cdot \int_{-\infty}^{\infty} x p_{\xi}(x) dx = c \cdot \mathrm{M}\{\xi\}.$$

3. $\mathrm{M}\{\xi \pm \eta\} = \mathrm{M}\{\xi\} \pm \mathrm{M}\{\eta\}$.

Пусть (ξ, η) – двумерная случайная величина с плотностью вероятностей $p_{\xi\eta}(x, y)$. Тогда

$$\mathrm{M}\{\xi \pm \eta\} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x \pm y)p_{\xi\eta}(x, y)dxdy =$$

$$= \int\limits_{-\infty}^{\infty} xdx \int\limits_{-\infty}^{\infty} p_{\xi\eta}(x, y)dy \pm \int\limits_{-\infty}^{\infty} ydy \int\limits_{-\infty}^{\infty} p_{\xi\eta}(x, y)dx =$$

$$= \int\limits_{-\infty}^{\infty} xp_{\xi}(x)dx \pm \int\limits_{-\infty}^{\infty} yp_{\eta}(y)dy = \mathrm{M}\{\xi\} \pm \mathrm{M}\{\eta\}.$$

Следствие. Пусть a и b – константы. Тогда

$$\mathrm{M}\{a\xi + b\} = a\mathrm{M}\{\xi\} + b.$$

4. Если ξ и η – **независимые** случайные величины, то

$$\mathrm{M}\{\xi\cdot\eta\}=\mathrm{M}\{\xi\}\cdot\mathrm{M}\{\eta\}.$$

Действительно, так как ξ и η независимы, то $p_{\xi\eta}(x,y)=p_{\xi}(x)\cdot p_{\eta}(y)$ и мы имеем

$$\mathrm{M}\{\xi\eta\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyp_{\xi\eta}(x,y)dxdy=$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyp_{\xi}(x)p_{\eta}(y)dxdy=\int_{-\infty}^{\infty}xp_{\xi}(x)dx\cdot\int_{-\infty}^{\infty}yp_{\eta}(y)dy=\mathrm{M}\{\xi\}\mathrm{M}\{\eta\}.$$

Существуют и другие характеристики положения случайных величин, например, медиана $Me\{\xi\}$, а для непрерывных величин – также мода $Mo\{\xi\}$.

Медиана

Другой достаточно популярной характеристикой положения случайной величины является так называемая медиана; её обычно обозначают $\mathrm{Me}\{\xi\}$.

Она определяется следующим соотношением

$$P\{\xi < \mathrm{Me}\{\xi\}\} = P\{\xi > \mathrm{Me}\{\xi\}\}.$$

Для непрерывных случайных величин медиану можно определить соотношением

$$\int_{-\infty}^{\mathrm{Me}\{\xi\}} p_{\xi}(x)dx = \int_{\mathrm{Me}\{\xi\}}^{\infty} p_{\xi}(x)dx = \frac{1}{2},$$

или, что то же самое, $F_{\xi}(\mathrm{Me}\{\xi\}) = 1/2$.

Модой непрерывной случайной величины называется то значение аргумента плотности распределения, при котором плотность достигает максимума.

Полезной числовой характеристикой непрерывных случайных величин являются квантили.

Квантилем порядка p называется корень $x = x_p$ уравнения $F(x) = p$. Очевидно, что квантиль порядка 1/2 является медианой.

Задача 6. На стороне монеты стоит 0, на другой – 1. Бросаем три таких монеты. Чему равно среднее значение выпавшей суммы очков?

Задача 7. Найти мат. ожидание суммы числа очков, выпавших на двух игральных костях.

Итак, если Пусть дана случайная величина ξ. Тогда её математическое ожидание $\mathrm{M}\{\xi\}$ определяется следующим образом:

$$\mathrm{M}\{\xi\} = \begin{cases} \sum\limits_{i=1}^{\infty} x_i p_i & \text{для дискретной случайной величины,} \\ \int\limits_{-\infty}^{\infty} x p_{\xi}(x) dx & \text{для непрерывной случайной величины.} \end{cases}$$

В самом общем виде математическое ожидание записывается через интеграл Стилтьеса:

$$\mathrm{M}\{\xi\} = \int\limits_{-\infty}^{\infty} x dF_{\xi}(x).$$

Дисперсия

Дисперсия относится к так называемым **характеристикам разброса**. Она показывает, насколько отдельные значения случайной величины могут отличаться между собой.

Дисперсией случайной величины называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания.

Дисперсия определяется следующим образом

$$\mathrm{D}\{\xi\} = \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})^2\},$$

или, в явном виде

$$\mathrm{D}\{\xi\} = \begin{cases} \sum\limits_{i=1}^{\infty}(x_i - \mathrm{M}\{\xi\})^2 p_i & \text{для дискретной случайной величины,} \\ \int\limits_{-\infty}^{\infty}(x - \mathrm{M}\{\xi\})^2 p_\xi(x)dx & \text{для непрерывной случайной величины} \end{cases}$$

Величина дисперсии имеет размерность квадрата от размерности самой случайной величины, поэтому часто в качестве характеристики разброса применяют **среднеквадратическое отклонение**.

Величина

$$\sigma\{\xi\} = \sqrt{\mathrm{D}\{\xi\}}$$

называется **среднеквадратическим отклонением** (с.к.о.) случайной величины ξ.

Для неотрицательных случайных величин полезной характеристикой является **коэффициент вариации** $v(\xi) = \dfrac{\sigma\{\xi\}}{M\{\xi\}}$.

Основные свойства дисперсии

1. $\mathrm{D}\{\xi\} = \mathrm{M}\{\xi^2\} - (\mathrm{M}\{\xi\})^2$.

Эта формула является также рабочей формулой для вычисления дисперсии. Надо лишь иметь в виду, что

$$\mathrm{M}\{\xi^2\} = \begin{cases} \sum\limits_{i=1}^{\infty} x_i^2 p_i & \text{для дискретной случайной величины,} \\ \int\limits_{-\infty}^{\infty} x^2 p_\xi(x) dx & \text{для непрерывной случайной величины.} \end{cases}$$

Дадим теперь вывод этой формулы. Имеем

$$(\xi - \mathrm{M}\{\xi\})^2 = \xi^2 - 2\xi\,\mathrm{M}\{\xi\} + (\mathrm{M}\{\xi\})^2.$$

Используя свойства математического ожидания, получаем

$$\mathrm{D}\{\xi\} = \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})^2\} = \mathrm{M}\{\xi^2\} - 2\,\mathrm{M}\{\xi\}\mathrm{M}\{\xi\} + \left(\mathrm{M}\{\xi\}\right)^2 =$$

$$= \mathrm{M}\{\xi^2\} - \left(\mathrm{M}\{\xi\}\right)^2.$$

2. Пусть c – константа. Тогда

$$\mathrm{D}\{c\} = 0; \quad \sigma\{c\} = 0.$$

Действительно, $\mathrm{M}\{c\} = c$, $\mathrm{M}\{c^2\} = c^2$, и поэтому

$$\mathrm{D}\{c\} = \mathrm{M}\{c^2\} - \mathrm{M}^2\{c\} = c^2 - c^2 = 0.$$

3. $D\{c\xi\} = c^2\, D\{\xi\}; \quad \sigma\{c\xi\} = |c|\,\sigma\{\xi\}$

Имеем

$$M\{c\xi\} = c \cdot M\{\xi\}; \qquad M\{c^2\xi^2\} = c^2 \cdot M\{\xi^2\}.$$

Поэтому

$$D\{c\xi\} = M\{c^2\xi^2\} - M^2\{c\xi\} = c^2\, M\{\xi^2\} - c^2\, M^2\{\xi\} =$$

$$= c^2\left(M\{\xi^2\} - M^2\{\xi\}\right) = c^2\, D\{\xi\}.$$

Извлекая из этого соотношения квадратный корень, получим

$$\sigma\{c\xi\} = \sqrt{D\{c\xi\}} = \sqrt{c^2\, D\{\xi\}} = |c|\,\sqrt{D\{\xi\}} = |c|\,\sigma\{\xi\}.$$

4. Если ξ и η **независимы**, то

$$D\{\xi \pm \eta\} = D\{\xi\} + D\{\eta\}.$$

Действительно, $M\{\xi \pm \eta\} = M\{\xi\} \pm M\{\eta\}$. Далее, так как

$$(\xi \pm \eta)^2 = \xi^2 \pm 2\xi\eta + \eta^2,$$

и величины ξ и η независимы, то

$$M\{(\xi \pm \eta)^2\} = M\{\xi^2\} \pm 2M\{\xi\}M\{\eta\} + M\{\eta^2\}.$$

Теперь имеем

$$D\{\xi \pm \eta\} = M\{(\xi \pm \eta)^2\} - M^2\{\xi \pm \eta\} =$$

$$= M\{\xi^2\} \pm 2M\{\xi\}M\{\eta\} + M\{\eta^2\} - \left(M^2\{\xi\} \pm 2M\{\xi\}M\{\eta\} + M^2\{\eta\}\right) =$$

$$= \left(M\{\xi^2\} - M^2\{\xi\}\right) + \left(M\{\eta^2\} - M^2\{\eta\}\right) = D\{\xi\} + D\{\eta\}.$$

Извлекая квадратный корень, получим

$$\sigma\{\xi \pm \eta\} = \sqrt{\sigma^2\{\xi\} + \sigma^2\{\eta\}}.$$

Задача 10. Подброшены две монеты. Случайная величина X – количество выпавших гербов. Найти математическое ожидание, дисперсию, среднее квадратическое отклонение.

Ряд распределения случайной величины X имеет вид

Значения x_i	0	1	2
Вероятность значения p_i	$1/4$	$1/2$	$1/4$

Математическое ожидание:

$M(X) = 0 \cdot 0.25 + 1 \cdot 0.5 + 2 \cdot 0.25 = 1$, т.е. при многократном проведении этого опыта среднее значение количества выпавших гербов будет близко к единице.

Дисперсия: $D(X) = M(X^2) - (M(X))^2 = 1.5 - 1 = 0.5$.

Среднее квадратичное отклонение по определению $\sigma(X) = \sqrt{D(X)} = \sqrt{0.5} = 0.707$.

Задача 11. По конвейеру вперемешку движутся карандаши, изготовленные из кедра (90%) или красного дерева (10%), а автомат раскладывает их по 5 штук в коробки. Пусть X – число карандашей из кедра в коробке. Составить для сл. величины X ряд распределения, найти среднее и дисперсию.

Решение. Случайная величина X подчиняется биномиальному распределению с параметрами $p = 0{,}9,\ q = 0{,}1,\ n = 5$. Испытание – выбор автоматом случайного карандаша. Успех – кедровая древесина карандаша.

$P_5(0) = C_5^0 \cdot p^0 \cdot q^5 = 0{,}00001,\ P_5(3) = C_5^3 \cdot p^3 \cdot q^2 = 0{,}0729$ и т.д.

Ряд распределения

X	0	1	2	3	4	5
p	0,00001	0,00045	0,0081	0,0729	0,325	0,59

$MX = n \cdot p = 5 \cdot 0{,}9 = 4{,}5,\ DX = n \cdot p \cdot q = 5 \cdot 0{,}9 \cdot 0{,}1 = 0{,}45$.

Моменты

И математическое ожидание и дисперсия являются представителями большой группы характеристик, получивших общее название «моменты случайных величин».

Начальным моментом *k*-го порядка случайной величины ξ называют величину

$$m_k(\xi) = \mathrm{M}\{\xi^k\} = \begin{cases} \sum\limits_{i=1}^{\infty} x_i^k p_i & \text{для дискретной случайной величины,} \\ \int\limits_{-\infty}^{\infty} x^k p_\xi(x)dx & \text{для непрерывной случайной величины} \end{cases}$$

В частности, $\mathrm{M}\{\xi\} = m_1(\xi)$.

Центральным моментом *k*-го порядка случайной величины ξ называют величину

$$\mu_k(\xi) = \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})^k\} =$$

$$= \begin{cases} \sum_{i=1}^{\infty} (x_i - \mathrm{M}\{\xi\})^k p_i & \text{для дискретной случайной величины,} \\ \int_{-\infty}^{\infty} (x - \mathrm{M}\{\xi\})^k p_\xi(x) dx & \text{для непрерывной случайной величины.} \end{cases}$$

В частности $M\xi = m_1, D\xi = \mu_2$.

Что касается $\mu_1(\xi)$, то

$$\mu_1(\xi) = \mathrm{M}\{\xi - \mathrm{M}\{\xi\}\} = \mathrm{M}\{\xi\} - \mathrm{M}\{\xi\} = 0.$$

На практике приходится встречаться еще с $\mu_3(\xi)$ и $\mu_4(\xi)$.

Центральный момент третьего порядка $\mu_3(\xi)$ определяет величину

$$Sk(\xi) = \frac{\mu_3(\xi)}{\sigma^3\{\xi\}},$$

которая называется **коэффициентом скошенности** или **коэффициентом асимметрии** случайной величины ξ. Она определяет вид плотности вероятностей величины ξ. Если $Sk(\xi) = 0$, то $p_\xi(x)$ симметрична около точки $x = \mathrm{M}\{\xi\}$ (см. рис. 1). Если $Sk(\xi) > 0$, то $p_\xi(x)$ имеет затянутый «хвост» справа (рис. 2); если же $Sk(\xi) < 0$, то $p_\xi(x)$ имеет затянутый «хвост» слева (рис. 3).

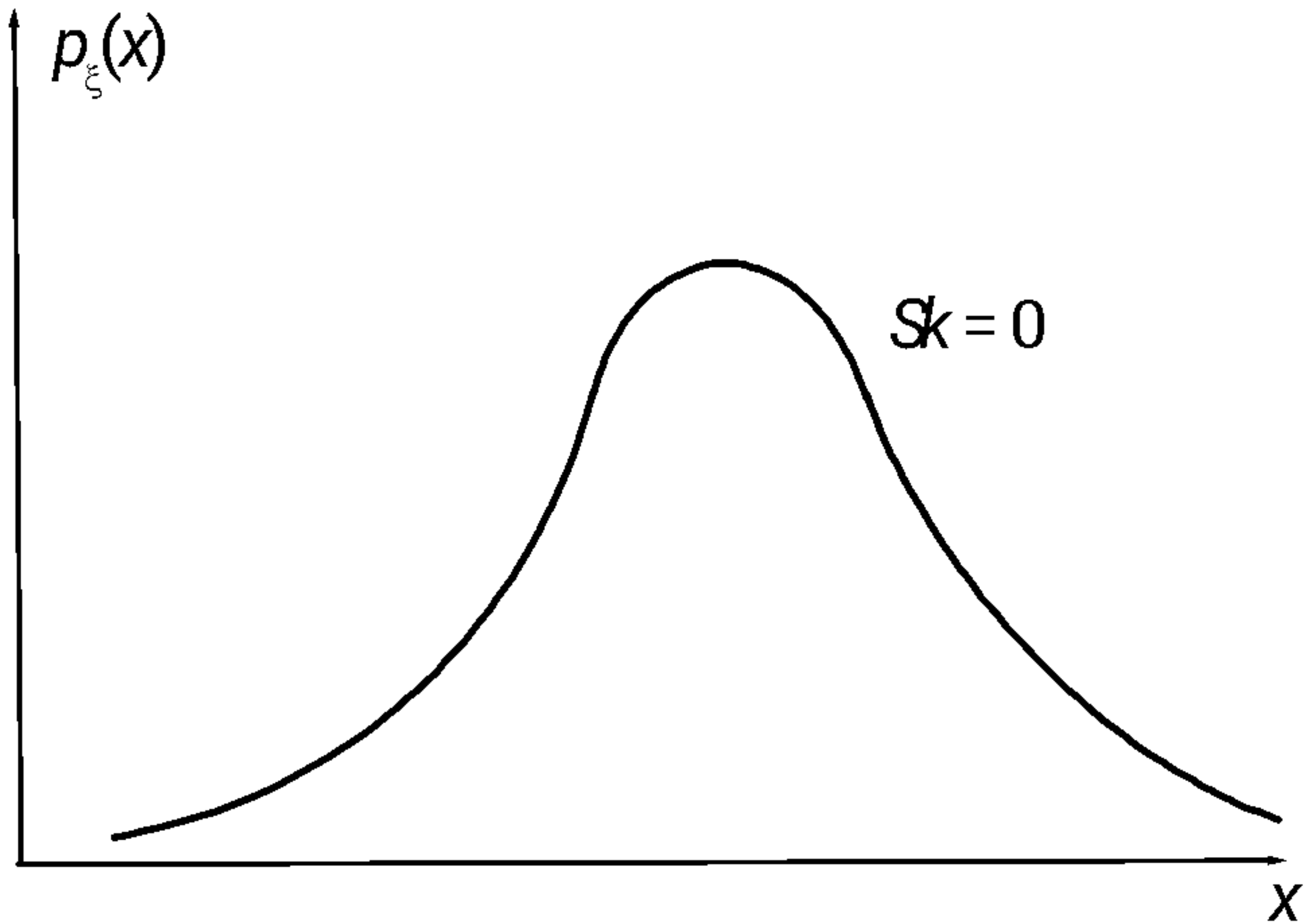


Рис.1

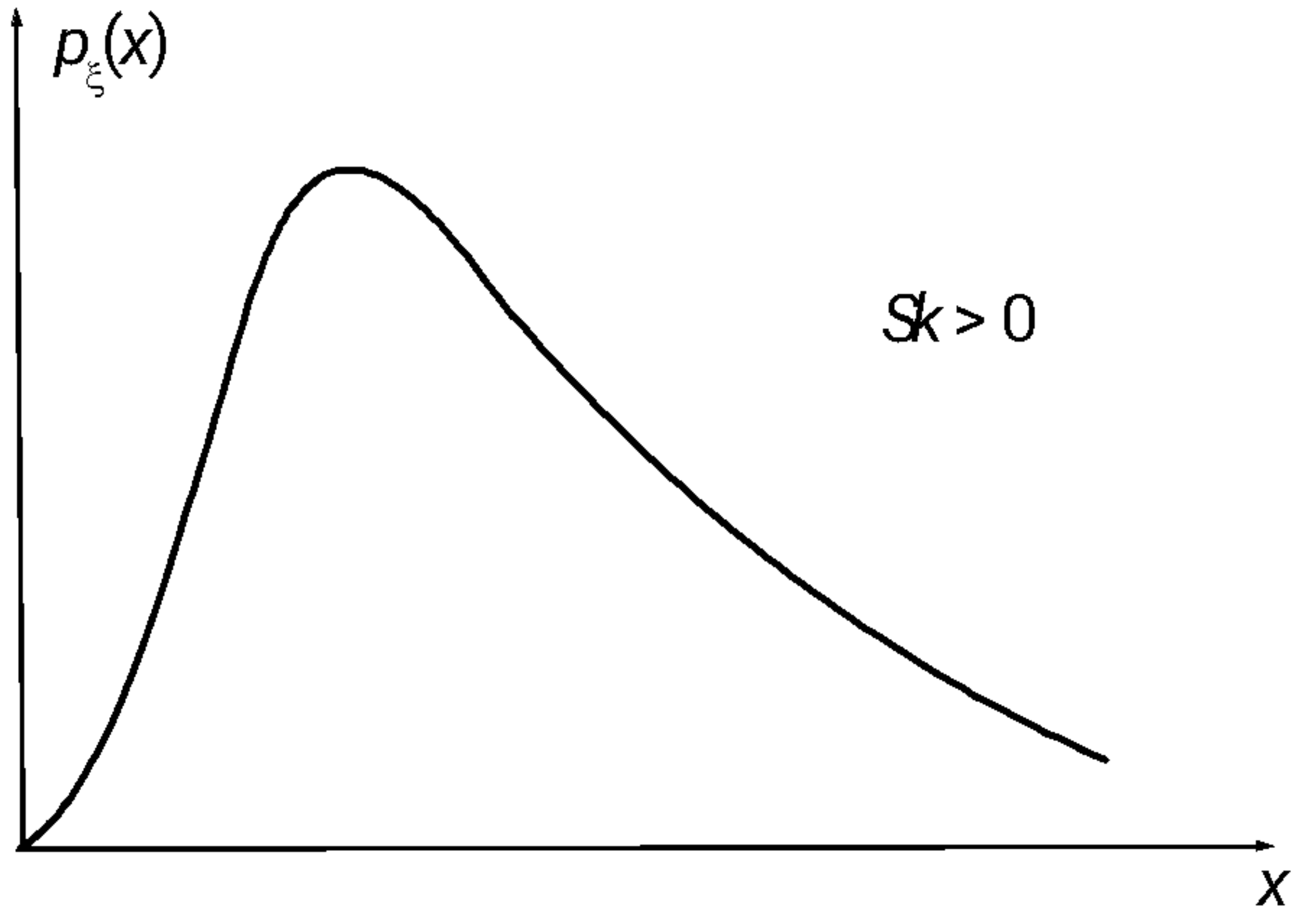


Рис.2

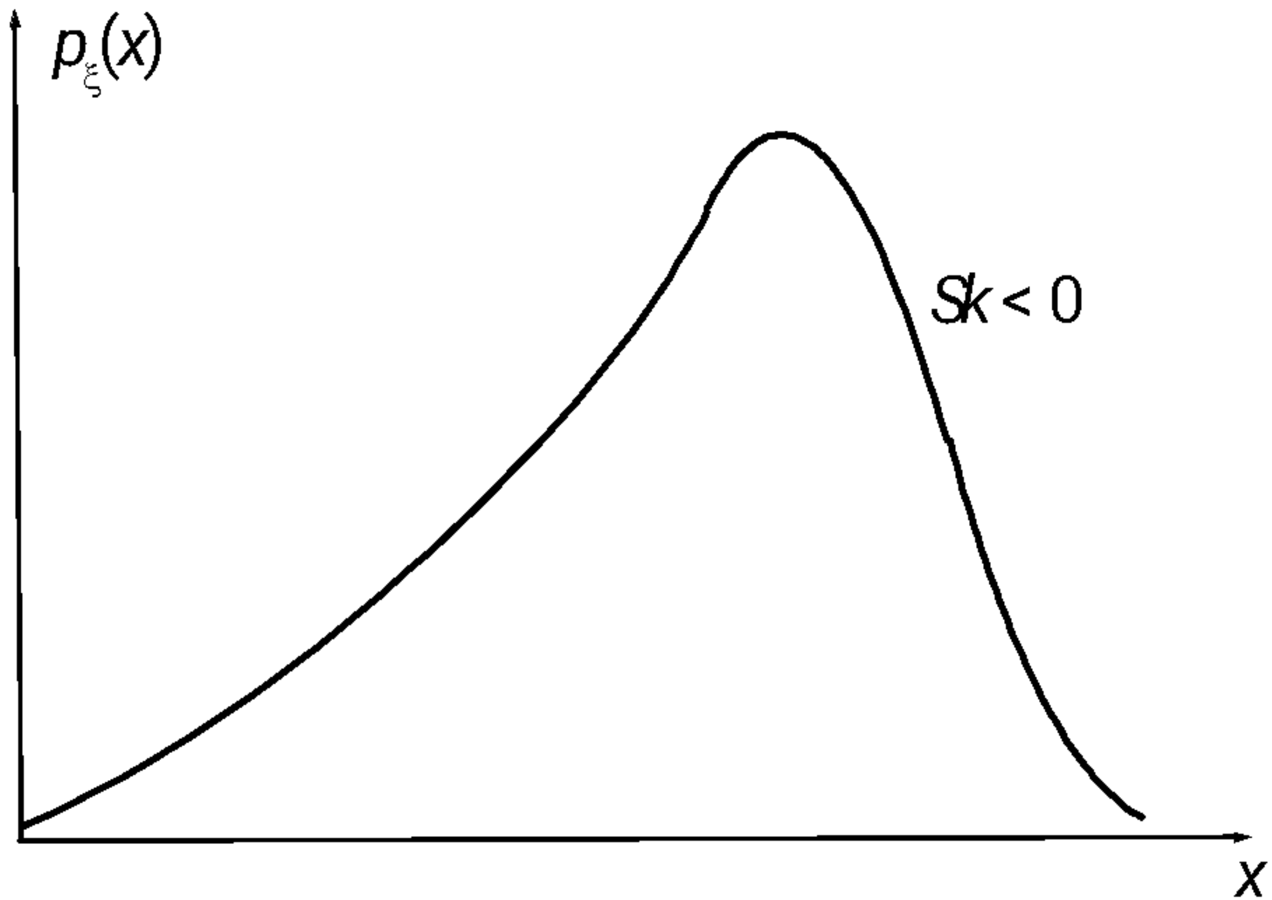


Рис. 3

Центральный момент четвёртого порядка $\mu_4(\xi)$ определяет величину

$$Ex(\xi) = \frac{\mu_4(\xi)}{\sigma^4\{\xi\}} - 3,$$

которую называют **эксцессом** или **коэффициентом островершинности** величины ξ. Дело в том, что в теории вероятностей есть некоторое стандартное распределение, называемое **нормальным**, о котором будет рассказано ниже. Если $Ex(\xi) > 0$, то наша $p_\xi(x)$ имеет более острую вершину, чем нормальное распределение (см. рис. 4). Если же $Ex(\xi) < 0$, то $p_\xi(x)$ имеет более плоскую вершину, чем нормальное распределение (см. рис. 5).

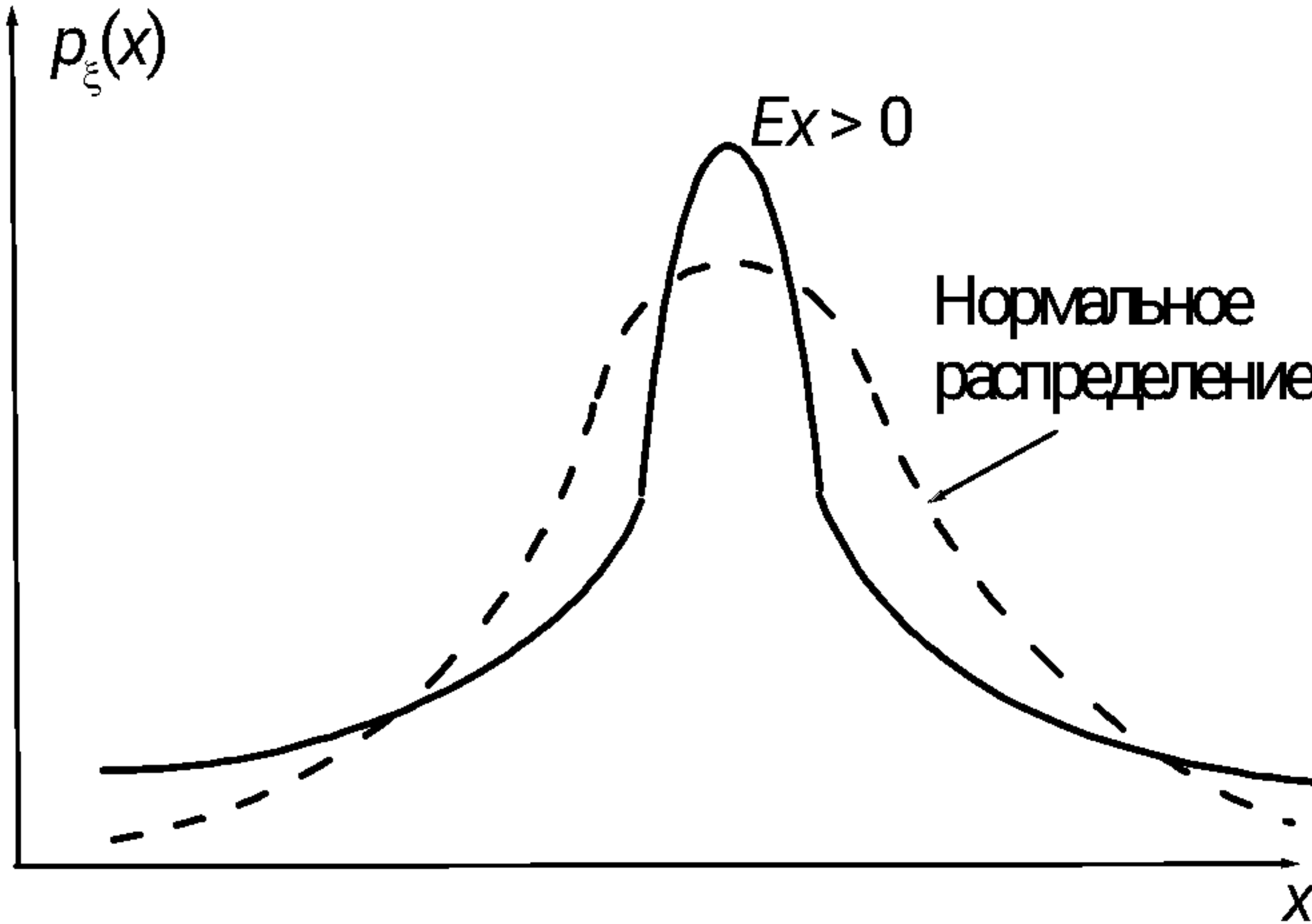


Рис. 4

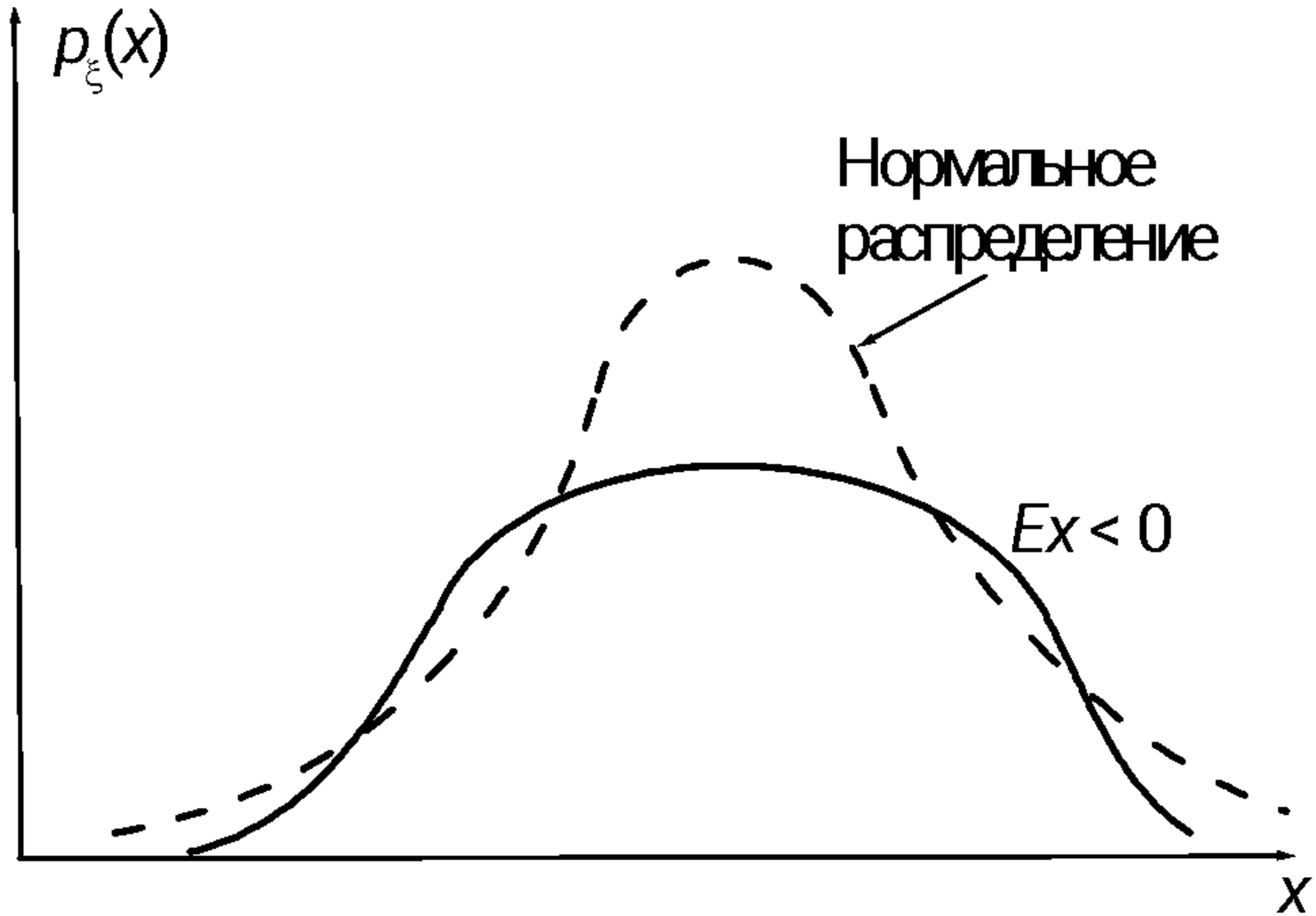


Рис.5

Между начальными и центральными моментами существуют соотношения:

$$\mu_k = M\left\{(\xi - m_1)^k\right\} = M\left\{\sum_{i=0}^{k} C_k^i \xi^i (-m_1)^{k-i}\right\} = \sum_{i=0}^{k} (-1)^{k-i} C_k^i m_i m_1^{k-i}$$

то есть центральные моменты k–го порядка μ_k определяются всеми начальными моментами m_i порядка $i, \forall\, 0 \le i \le k$. Здесь $m_0 = 1$.

С другой стороны, имеем

$$m_k = M\left\{\xi^k\right\} = M\left\{\left[(\xi - m_1) + m_1\right]^k\right\} = M\left\{\sum_{i=0}^{k} C_k^i (\xi - m_1)^i m_1^{k-i}\right\} = \sum_{i=0}^{k} C_k^i \mu_i m_1^{k-i}$$

т.е. начальные моменты k–го порядка m_k определяются начальными моментами первого порядка m_1 и всеми центральными моментами порядка $i, \forall\, 0 \le i \le k$. Здесь $\mu_0 = 1, \mu_1 = 0$.

На практике моменты порядка выше четвертого применяются достаточно редко

Кривые регрессии

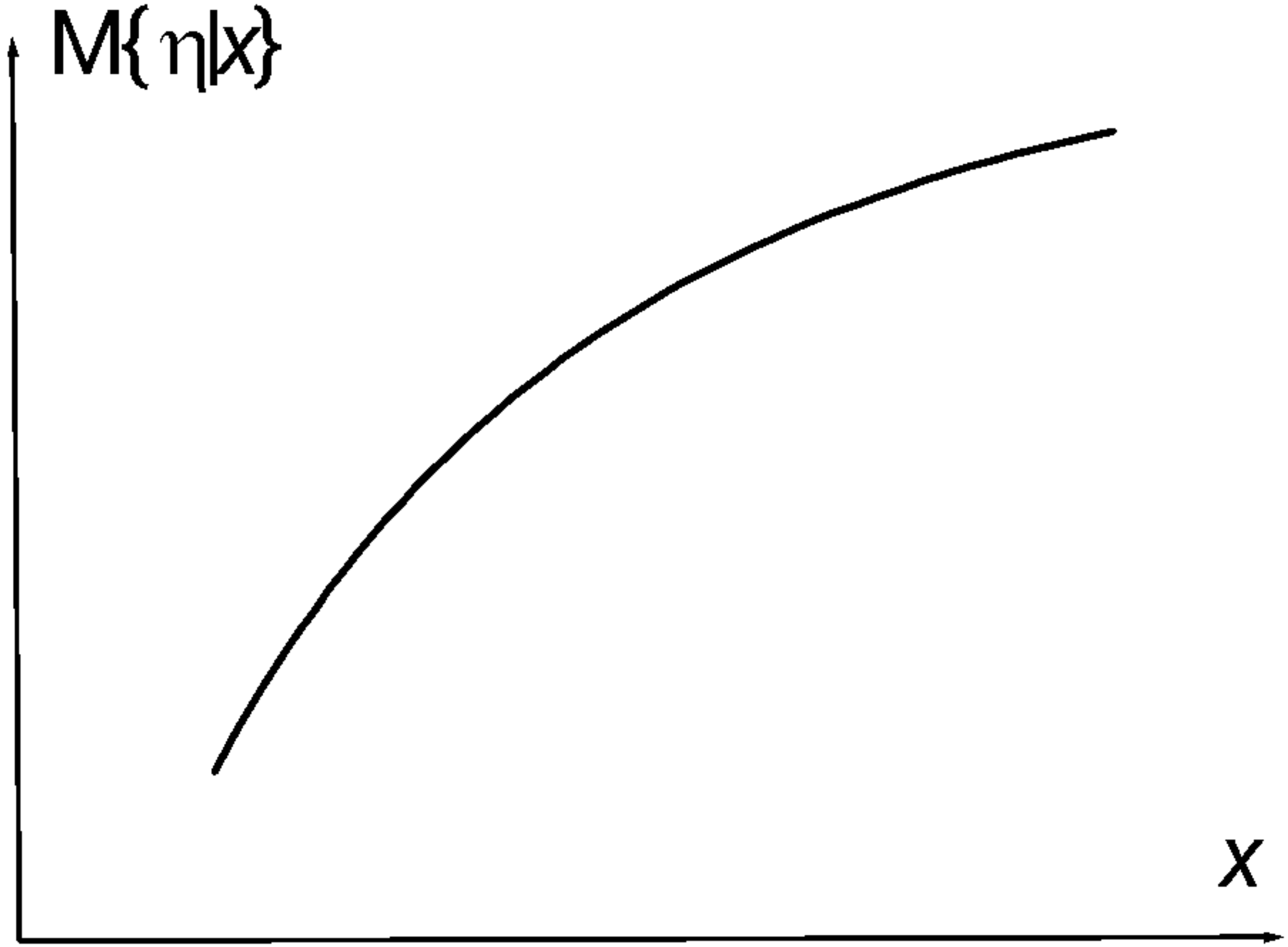
Последняя группа характеристик – это характеристики связи, зависимости случайных величин друг от друга.

Наиболее полно зависимость между случайными величинами ξ и η описывается условными функциями распределения $F_{\xi|\eta}(x,y)$ и $F_{\eta|\xi}(y,x)$ или условными плотностями вероятностей $p_{\xi|\eta}(x|y)$ и $p_{\eta|\xi}(y|x)$. Однако эти величины сложны для экспериментального измерения, и с ними трудно работать.

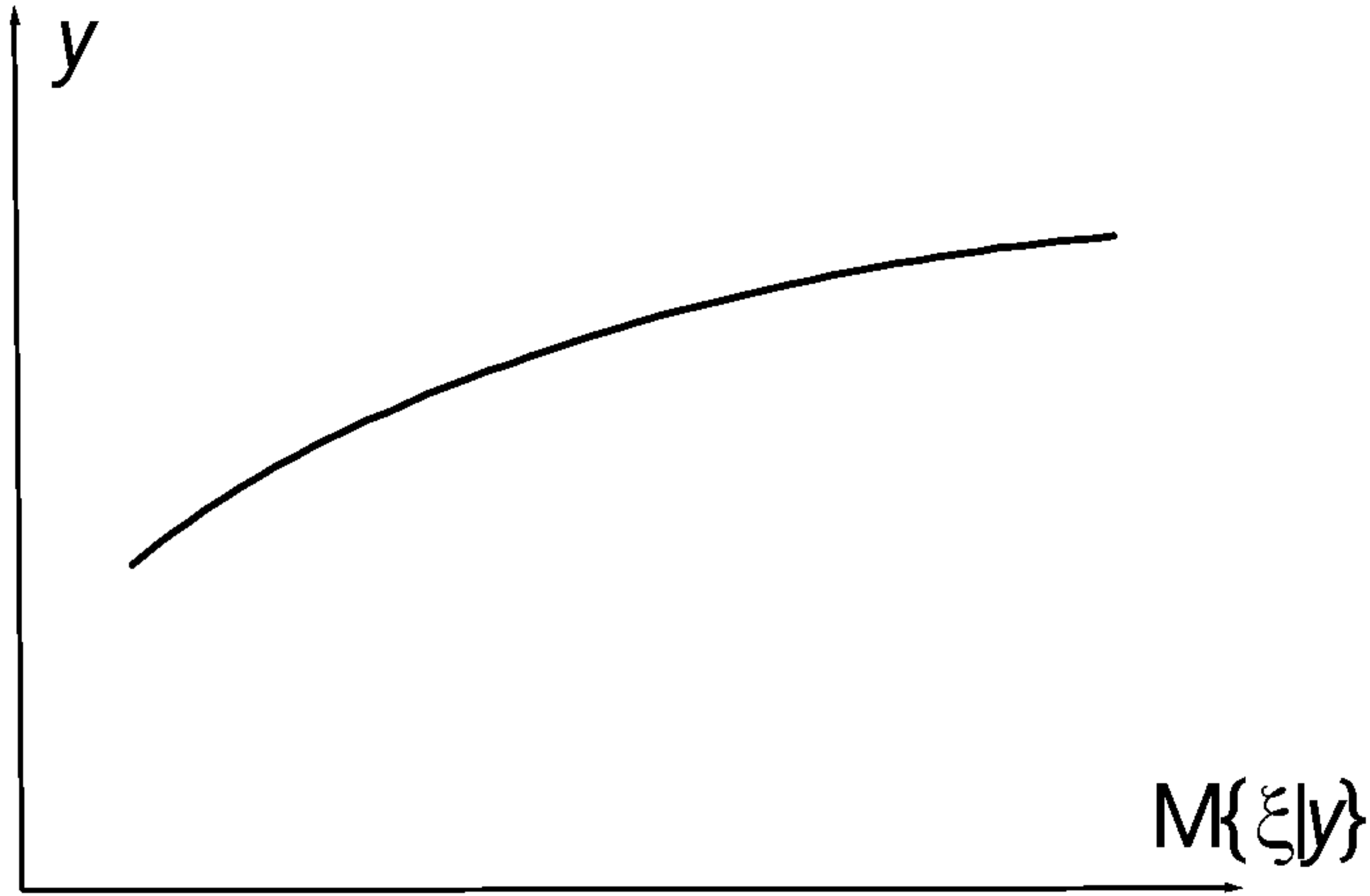
Вместо них рассмотрим так называемые условные математические ожидания (условные средние)

$$\mathrm{M}\{\eta\,|\,x\} = \int_{-\infty}^{\infty} y p_{\eta|\xi}(y\,|\,x)dy, \qquad \mathrm{M}\{\xi\,|\,y\} = \int_{-\infty}^{\infty} x p_{\xi|\eta}(x\,|\,y)dx.$$

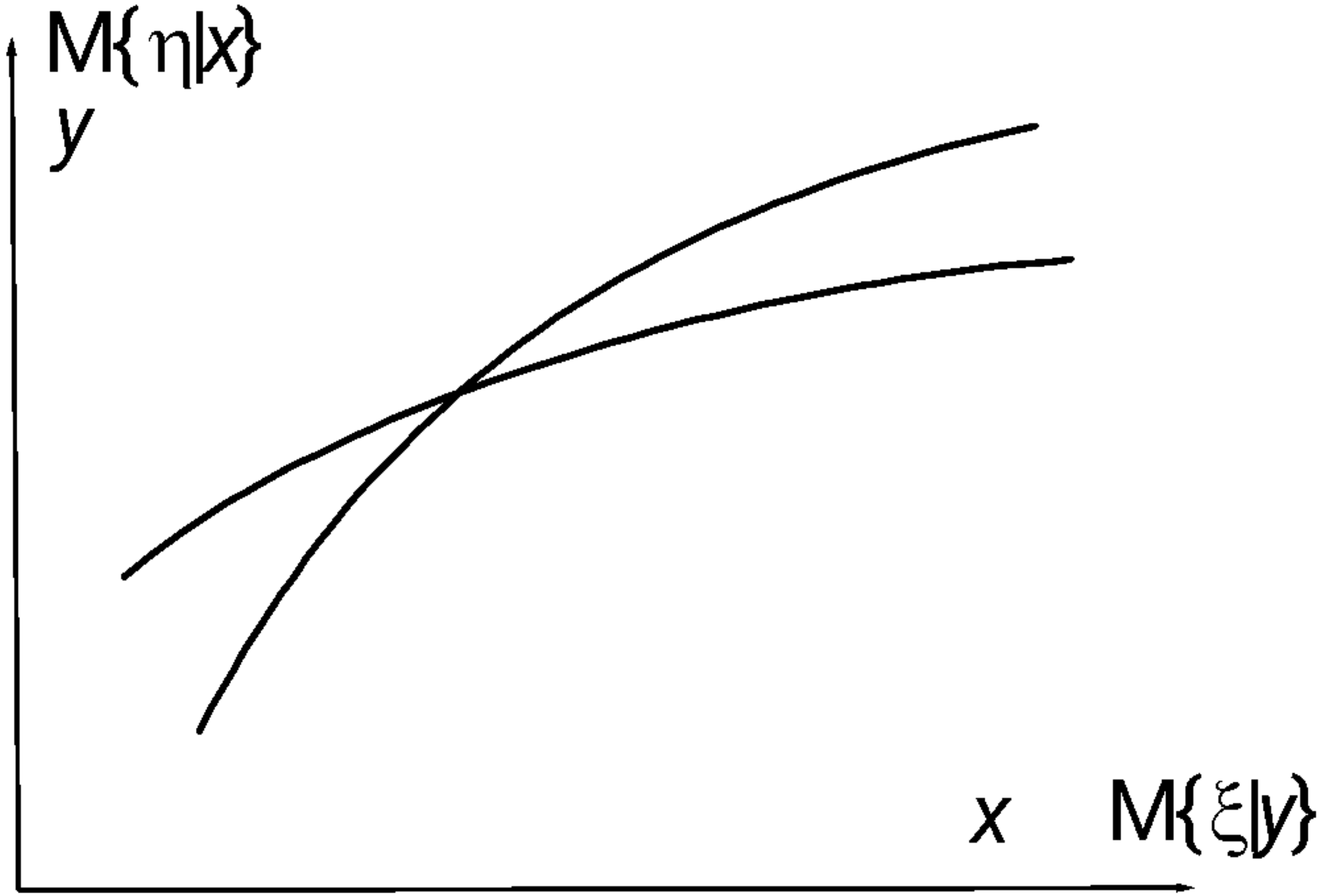
В отличие от M{ξ} и M{η} это уже не числа, а **функции**. Кривую $\mathrm{M}\{\eta|\,x\}$ как функцию от x называют **кривой регрессии η от ξ,** а $\mathrm{M}\{\xi|\,y\}$, являющейся функцией от y, называют **кривой регрессии** ξ от η.



Кривая регрессии η от ξ



Кривая регрессии ξ от η



Обе кривые регрессии на одном графике

Ковариация и коэффициент корреляции

Описание зависимости случайных величин кривыми регрессии также является достаточно сложным, поэтому возникает необходимость дальнейшего упрощения этого описания какой-либо числовой характеристикой.

Пусть дана двумерная случайная величина (ξ, η) с плотностью вероятностей $p_{\xi\eta}(x, y)$. **Ковариацией** величин ξ и η называется величина

$$\operatorname{cov}(\xi,\eta) = \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})(\eta - \mathrm{M}\{\eta\})\} =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mathrm{M}\{\xi\})(y - \mathrm{M}\{\eta\})\, p_{\xi\eta}(x,y)\,dxdy.$$

Величина

$$r_{\xi\eta} = \frac{\operatorname{cov}(\xi,\eta)}{\sqrt{D\{\xi\}\cdot D\{\eta\}}} = \frac{\operatorname{cov}(\xi,\eta)}{\sigma\{\xi\}\cdot\sigma\{\eta\}}$$

называется **коэффициентом корреляции** величин ξ и η.

Чтобы понять, что характеризует коэффициент корреляции, рассмотрим его свойства.

1. $-1 \le r_{\xi\eta} \le +1$.

Для доказательства рассмотрим комбинацию

$$\frac{\xi - \mathrm{M}\{\xi\}}{\sqrt{\mathrm{D}\{\xi\}}} \pm \frac{\eta - \mathrm{M}\{\eta\}}{\sqrt{\mathrm{D}\{\eta\}}}.$$

Так как квадрат любой величины неотрицателен, то

$$\left(\frac{\xi - \mathrm{M}\{\xi\}}{\sqrt{\mathrm{D}\{\xi\}}} \pm \frac{\eta - \mathrm{M}\{\eta\}}{\sqrt{\mathrm{D}\{\eta\}}}\right)^2 \ge 0,$$

и поэтому

$$\mathrm{M}\left\{\left(\frac{\xi - \mathrm{M}\{\xi\}}{\sqrt{\mathrm{D}\{\xi\}}} \pm \frac{\eta - \mathrm{M}\{\eta\}}{\sqrt{\mathrm{D}\{\eta\}}}\right)^2\right\} \ge 0.$$

Раскрывая квадрат, получим

$$\mathrm{M}\left\{\frac{(\xi-\mathrm{M}\{\xi\})^2}{\mathrm{D}\{\xi\}}\pm 2\frac{(\xi-\mathrm{M}\{\xi\})(\eta-\mathrm{M}\{\eta\})}{\sqrt{\mathrm{D}\{\xi\}\mathrm{D}\{\eta\}}}+\frac{(\eta-\mathrm{M}\{\eta\})^2}{\mathrm{D}\{\eta\}}\right\}\geq 0.$$

Пользуясь свойствами математического ожидания, будем иметь

$$\frac{\mathrm{D}\{\xi\}}{\mathrm{D}\{\xi\}}\pm 2\frac{\mathrm{cov}(\xi,\eta)}{\sqrt{\mathrm{D}\{\xi\}\mathrm{D}\{\eta\}}}+\frac{\mathrm{D}\{\eta\}}{\mathrm{D}\{\eta\}}\geq 0,$$

или

$$1\pm 2r_{\xi\eta}+1\geq 0,\quad \text{или } 1\pm r_{\xi\eta}\geq 0.$$

Беря в этом выражении знак «+», получим, что $r_{\xi\eta}\geq -1$, а при знаке «–» получим, что $r_{\xi\eta}\leq 1$. Объединяя все это вместе, получим, что $-1\leq r_{\xi\eta}\leq +1$ или

$|\,r_{\xi\eta}\,|\leq 1.$

2. Если случайные величины ξ и η **независимы,** то $r_{\xi\eta} = 0$.

Действительно, для независимых ξ и η имеем

$$\operatorname{cov}(\xi, \eta) = \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})(\eta - \mathrm{M}\{\eta\})\} =$$

$$= \mathrm{M}\{\xi - \mathrm{M}\{\xi\}\} \cdot \mathrm{M}\{\eta - \mathrm{M}\{\eta\}\} = \big(\mathrm{M}\{\xi\} - \mathrm{M}\{\xi\}\big) \cdot \big(\mathrm{M}\{\eta\} - \mathrm{M}\{\eta\}\big) = 0,$$

и поэтому $r_{\xi\eta} = \dfrac{\operatorname{cov}(\xi, \eta)}{\sigma\{\xi\} \cdot \sigma\{\eta\}} = 0$.

3. Если $\eta = a\xi + b$, то $r_{\xi\eta} = \pm 1$.

Действительно, в этом случае

$$\mathrm{M}\{\eta\} = a\mathrm{M}\{\xi\} + b, \quad \mathrm{D}\{\eta\} = a^2\,\mathrm{D}\{\xi\},$$

$$\operatorname{cov}(\xi,\eta) = \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})(\eta - \mathrm{M}\{\eta\})\} =$$

$$= \mathrm{M}\{(\xi - \mathrm{M}\{\xi\})(a\xi + b - a\mathrm{M}\{\xi\} - b)\} =$$

$$= a\mathrm{M}\{(\xi - \mathrm{M}\{\xi\})^2\} = a\mathrm{D}\{\xi\}$$

$$D\{\eta\} = M\left\{(\eta - M\eta)^2\right\} = M\left\{(a\xi + b - aM\xi - b)^2\right\} = a^2 D\xi$$

и поэтому

$$r_{\xi\eta} = \frac{a\mathrm{D}\{\xi\}}{\sqrt{\mathrm{D}\{\xi\}\cdot a^2\,\mathrm{D}\{\xi\}}} = \frac{a}{|a|} = \begin{cases} +1, & \text{если } a > 0, \\ -1, & \text{если } a < 0. \end{cases}$$

Оказывается, что можно доказать и обратное – если $r_{\xi\eta} = \pm 1$, то существуют такие a и b, что $\eta = a\xi + b$.

4. Если $r_{\xi\eta} = \pm 1$, то в среднем квадратическом выполняется равенство $\eta = a\xi + b$, где a и b – неслучайные числа.

В свойстве 1. было показано, что $M\left\{\left(\frac{\xi - M\xi}{\sigma\{\xi\}} \pm \frac{\eta - M\eta}{\sigma\{\eta\}}\right)^2\right\} = 2\left(1 \pm r_{\xi\eta}\right)$,

поэтому, если $r = 1$, то выбирая в этом равенстве знак –, а если $r = -1$, то, выбирая знак +, получим, что в среднем квадратическом выполняется равенство $\frac{\eta - M\eta}{\sigma\{\eta\}} - r\frac{\xi - M\xi}{\sigma\{\xi\}} = 0$, откуда следует, что

$\frac{\eta - M\eta}{\sigma\{\eta\}} = r\frac{\xi - M\xi}{\sigma\{\xi\}}$, тогда

$$\eta = r\frac{\xi - M\xi}{\sigma\{\xi\}}\sigma\{\eta\} + M\eta = \left\{r\frac{\sigma\{\eta\}}{\sigma\{\xi\}}\right\}\xi + \left\{M\eta - r\frac{\sigma\{\eta\}}{\sigma\{\xi\}}M\xi\right\} = a\xi + b.$$

5. Можно привести пример такой нелинейной зависимости случайных величин ξ и η, для которых $r_{\xi\eta} = 0$.

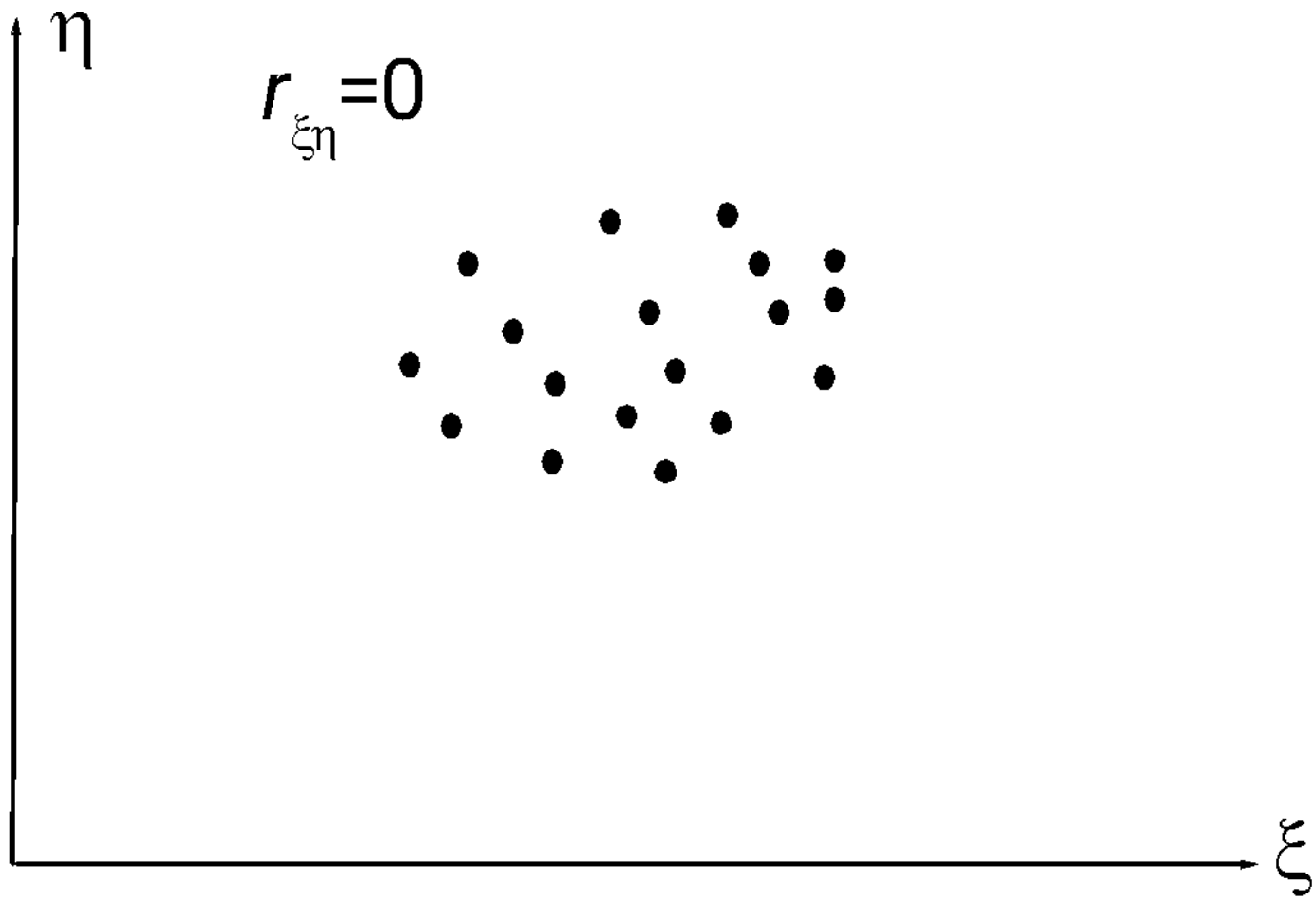
Пусть ξ такая случайная величина, для которой $p_{\xi}(x)$ есть чётная функция, так что $\mathrm{M}\{\xi\} = 0$ и $\mathrm{M}\{\xi^3\} = 0$. Пусть далее $\eta = \xi^2$, то есть η связано с ξ функциональной зависимостью. Тогда $\mathrm{M}\{\eta\} = \mathrm{M}\{\xi^2\}$ и мы получаем

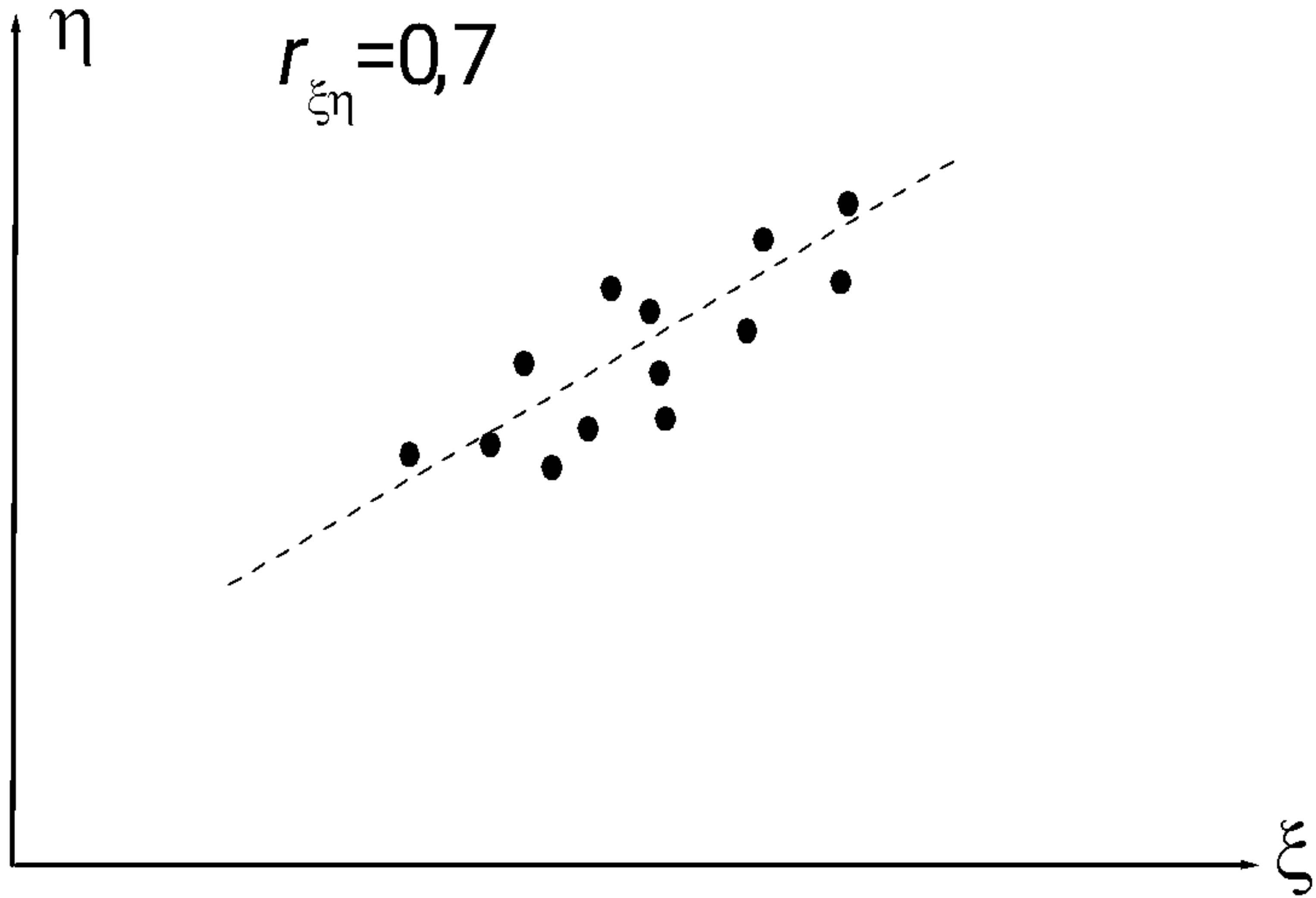
$$\mathrm{cov}(\xi, \eta) = \mathrm{M}\{\xi \cdot (\xi^2 - \mathrm{M}\{\xi^2\})\} =$$

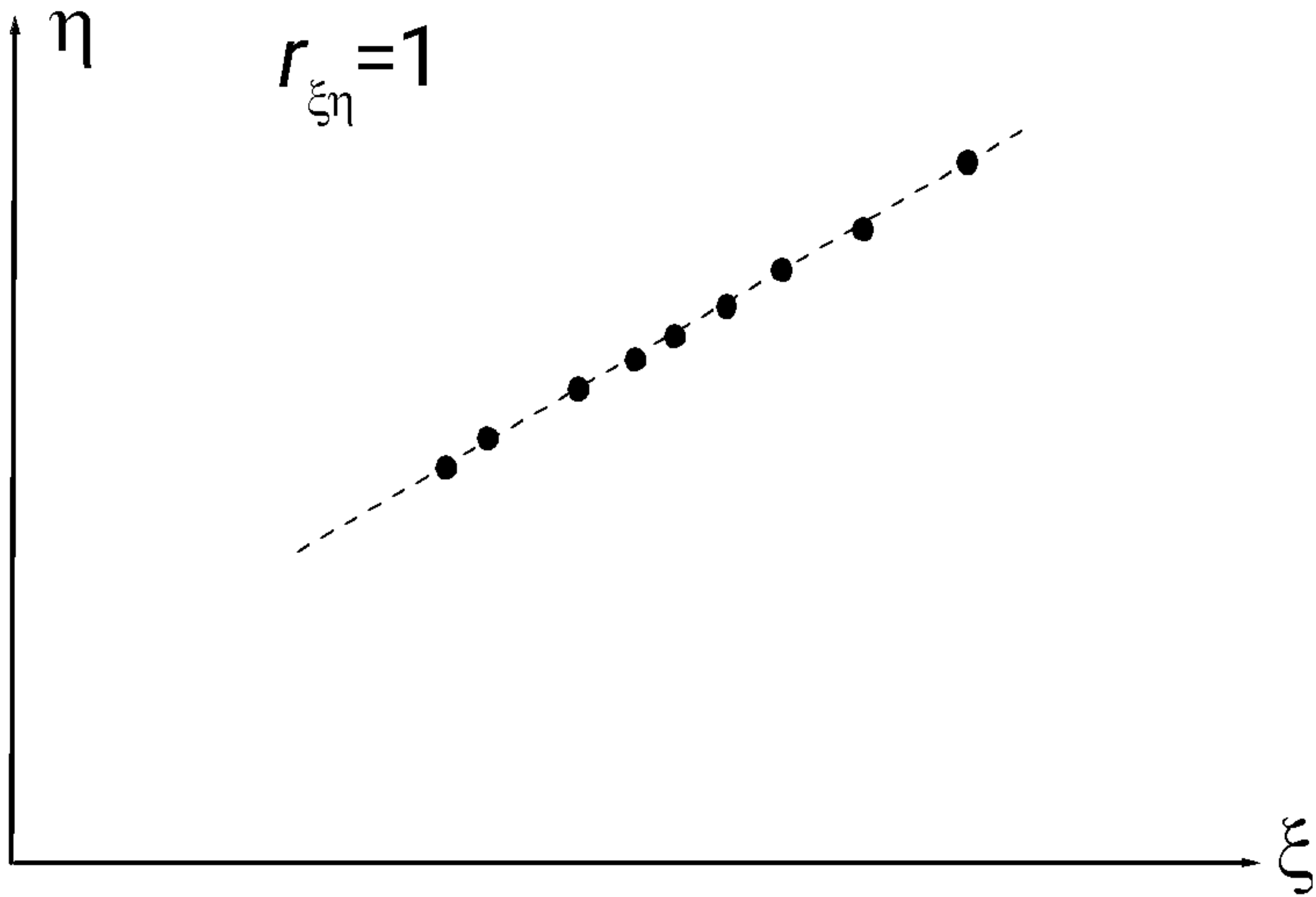
$$= \mathrm{M}\{\xi^3\} - \mathrm{M}\{\xi^2\} \cdot \mathrm{M}\{\xi\} = 0$$

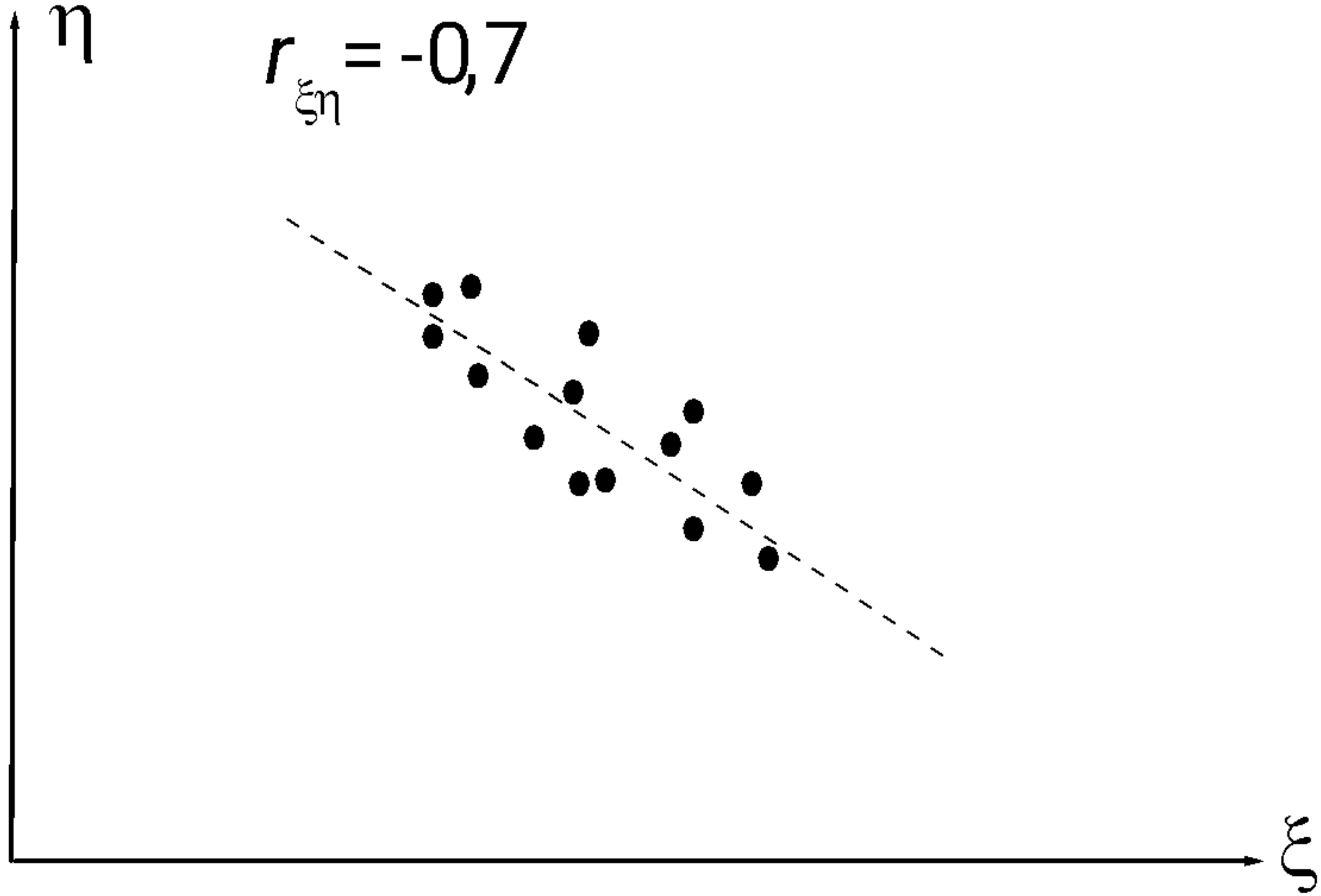
и поэтому в данной ситуации $r_{\xi\eta} = 0$.

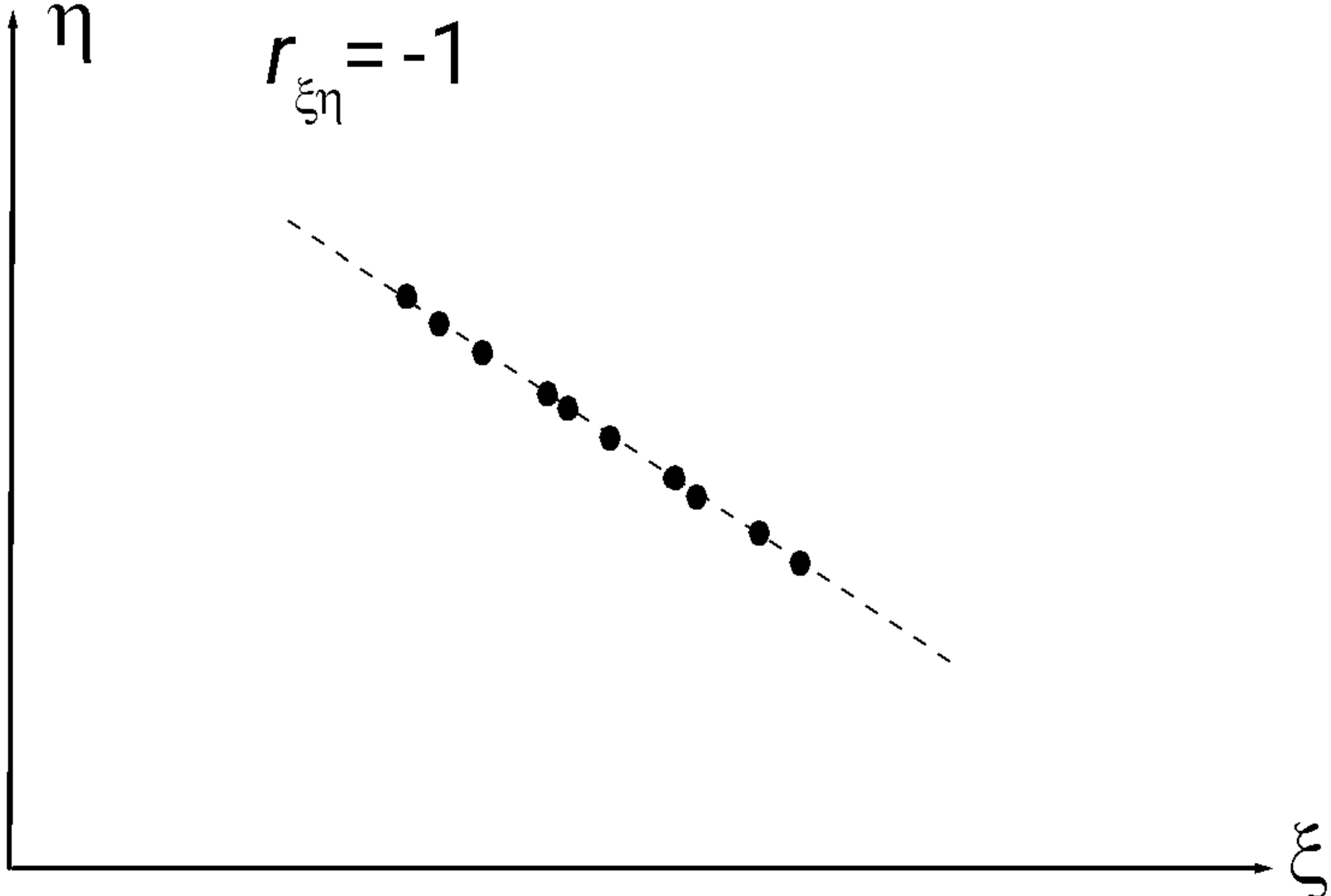
Подводя итог этим свойствам можно сказать, что коэффициент корреляции, конечно, отражает зависимость между случайными величинами, но **не всякую зависимость**. Как говорят, коэффициент корреляции является мерой **линейной зависимости** случайных величин друг от друга, так как с увеличением $|r_{\xi\eta}|$ зависимость между ними все более приближается к линейной.


η
$r_{\xi\eta}=0$
ξ


η
$r_{\xi\eta}$=0,7
ξ


η
$r_{\xi\eta}=1$
ξ


η
$r_{\xi\eta} = -0,7$
ξ


η
$r_{\xi\eta} = -1$
ξ