Правило 3-х σ (трех “сигм”).

Пусть имеется нормально распределённая случайная величина ξ с математическим ожиданием, равным a и дисперсией σ^2. Определим вероятность попадания ξ в интервал $(a - 3\sigma; a + 3\sigma)$, то есть вероятность того, что ξ принимает значения, отличающиеся от математического ожидания не более, чем на три среднеквадратических отклонения.

$$P(a - 3\sigma < \xi < a + 3\sigma) = \Phi(3) - \Phi(-3) = 2\Phi(3)$$

По таблице находим $\Phi(3)=0{,}49865$, откуда следует, что $2\Phi(3)$ практически равняется единице. Таким образом, можно сделать важный вывод: нормальная случайная величина принимает значения, отклоняющиеся от ее математического ожидания не более чем на 3σ.

(Выбор числа 3 здесь условен и никак не обосновывается: можно было выбрать 2,8, 2,9 или 3,2 и получить тот же вероятностный результат. Учитывая, что $\Phi(2)=0{,}477$, можно было бы говорить и о правиле 2–х “сигм”.)

Совместное распределение двух случайных величин.

Пусть пространство элементарных исходов Ω случайного эксперимента таково, что каждому исходу ω_i^j ставиться в соответствие значение случайной величины ξ, равное x_i и значение случайной величины η, равное y^j.

Примеры:

1. Представим себе большую совокупность деталей, имеющих вид стержня. Случайный эксперимент заключается в случайном выборе одного стержня. Этот стержень имеет длину, которую будем обозначать ξ и толщину—η (можно указать другие параметры—объем, вес, чистота обработки, выраженная в стандартных единицах).
2. Если результат эксперимента—случайный выбор какого–либо предприятия в данной области, то за ξ можно принимать объем производства отнесенный к количеству сотрудников, а за η—объем продукции, идущей на экспорт, тоже отнесенной к числу сотрудников.

В этом случае мы можем говорить о совместном распределении случайных величин ξ и η или о “двумерной” случайной величине.

Если ξ и η дискретны и принимают конечное число значений (ξ – n значений, а η – k значений), то закон совместного распределения случайных величин ξ и η можно задать, если каждой паре чисел x_i, y^j (где x_i принадлежит множеству значений ξ, а y^j—множеству значений η) поставить в соответствие вероятность p_i^j, равную вероятности события, объединяющего все исходы ω_i^j (и состоящего лишь из этих исходов), которые приводят к значениям

$$\xi = x_i;\ \eta = y^j.$$

Такой закон распределения можно задать в виде таблицы:

η / ξ	y^1	y^2	…	y^j	…	y^k	
x_1	p_1^1	p_1^2	…	p_1^j	…	p_1^k	P_1
…	…	…	…	…	…	…	…
x_i	p_i^1	p_i^2	…	p_i^j	…	p_i^k	P_i
…	…	…	…	…	…	…	…
x_n	p_n^1	p_n^2	…	p_n^j	…	p_n^k	P_n
	P^1	P^2	…	P^j	…	P^k	…

(*)

Очевидно $\sum_{i=1}^{n}\sum_{j=1}^{k} p_i^j = 1$

Если просуммировать все p_i^j в i–й строке, то получим

$$\sum_{j=1}^{k} p_i^j = P_i$$

вероятность того, что случайная величина ξ примет значение x_i. Аналогично, если просуммировать все p_i^j в j–м столбце, то получим

$$\sum_{i=1}^{n} p_i^j = P^j$$

вероятность того, что η принимает значение y^j.

Соответствие $x_i \to P_i$ $(i = 1,2,\ldots,n)$ определяет закон распределения ξ, также как соответствие $y^j \to P^j$ $(j = 1,2,\ldots,k)$ определяет закон распределения случайной величины η.

Очевидно $M_\xi = \sum_{i=1}^{n} x_i P_i$, $M_\eta = \sum_{j=1}^{k} y^j P^j$.

Раньше мы говорили, что случайные величины ξ и η независимы, если

$$p_i^j = P_i \cdot P^j \quad (i=1,2,\ldots,n;\ j=1,2,\ldots,k).$$

Если это не выполняется, то ξ и η зависимы.

В чем проявляется зависимость случайных величин ξ и η и как ее выявить из таблицы?

Рассмотрим столбец y^1. Каждому числу x_i поставим в соответствие число

$$p_{i/1} = \frac{p_i^1}{P_1} \qquad (1)$$

которое будем называть условной вероятностью $\xi = x_i$ при $\eta = y^1$. Обратите внимание на то, что это не вероятность P_i события $\xi = x_i$, и сравните формулу (1) с уже известной формулой условной вероятности $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$.

Соответствие

$$x_i \to p_{i/1},\ (i=1,2,\ldots,n)$$

будем называть условным распределением случайной величины ξ при $\eta = y^1$. Очевидно $\sum_{i=1}^{n} p_{i/1} = 1$.

Аналогичные условные законы распределения случайной величины ξ можно построить при всех остальных значениях η, равных y^2; y^3,…, y^n ,ставя в соответствие числу x_i условную вероятность $p_{i/j} = \frac{p_i^j}{P_j}$ $(\sum_{i=1}^{n} p_{i/j} = 1)$.

В таблице приведён условный закон распределения случайной величины ξ при $\eta = y^j$

ξ	x_1	x_2	…	x_i	…	x_n
$p_{i/j}$	$\frac{p_1^j}{P_j}$	$\frac{p_2^j}{P_j}$	…	$\frac{p_i^j}{P_j}$	…	$\frac{p_n^j}{P_j}$

Можно ввести понятие условного математического ожидания ξ при $\eta = y^j$

$$M(\xi / \eta = y_j) = \sum_{i=1}^{n} x_i \frac{p_i^j}{P^j} = \frac{1}{P^j} \sum_{i=1}^{n} x_i p_i^j$$

Заметим, что ξ и η равноценны. Можно ввести условное распределение η при $\xi = x_i$ соответствием

$$y^j \to \frac{p_i^j}{P_i} \quad (j = 1,2,\ldots,k)$$

Также можно ввести понятие условного математического ожидания случайной величины η при $\xi = x_i$:

$$M(\eta / \xi = x_i) = \sum_{j=1}^{k} y^j \frac{p_i^j}{P_i} = \frac{1}{P_i} \sum_{j=1}^{k} y^j p_i^j$$

Из определения следует, что если ξ и η независимы, то все условные законы распределения одинаковы и совпадают с законом распределения ξ (напоминаем, что закон распределения ξ определяется в таблице (*) первым и последним столбцом). При этом очевидно, совпадают все условные математические ожидания $M(\xi/\eta = y^j)$ при $j = 1,2,\ldots,k$, которые равны Mξ.

Если условные законы распределения ξ при различных значениях η различны, то говорят, что между ξ и η имеет место статистическая зависимость.

Пример I. Пусть закон совместного распределения двух случайных величин ξ и η задан следующей таблицей. Здесь, как говорилось ранее, первый и последний столбцы определяют закон распределения случайной величины ξ, а первая и последняя строки – закон распределения случайной величины η.

ξ \ η	1	2	3	
10	1/36	0	0	1/36
20	2/36	1/36	0	3/36
30	2/36	3/36	2/36	7/36
40	1/36	8/36	16/36	25/36
	6/36	12/36	18/36	

Полигоны условных распределений можно изобразить на трехмерном графике (рис. 1).

Здесь явно просматривается зависимость условного закона распределения ξ от величины η.

Пример II. (Уже встречавшийся).

Пусть даны две независимые случайные величины ξ и η с законами распределения

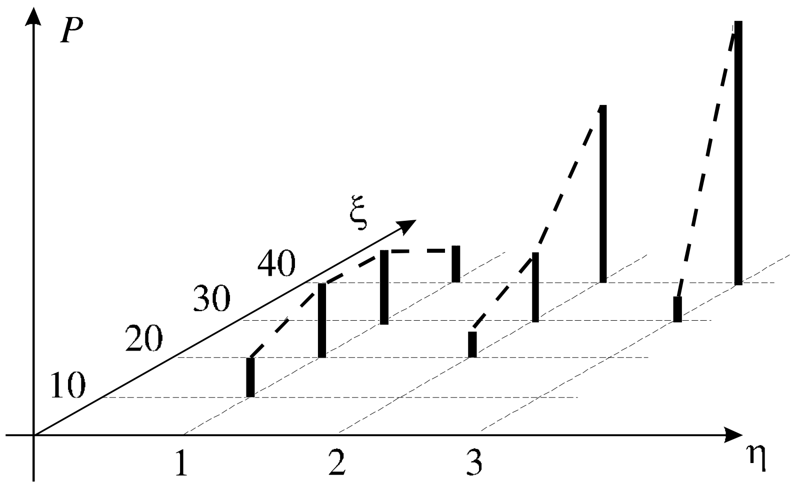

Рис. 1

ξ	0	1
P	1/3	2/3

η	1	2
P	3/4	1/4

Найдем законы распределений случайных величин $\alpha=\xi+\eta$ и $\beta=\xi*\eta$

α	1	2	3
P	3/12	7/12	2/12

β	0	1	2
P	4/12	6/12	2/12

Построим таблицу закона совместного распределения α и β.

α \ β	0	1	2	
1	3/12	0	0	3/12
2	1/12	6/12	0	7/12
3	0	0	2/12	2/12
	4/12	6/12	2/12	

Чтобы получить $\alpha=2$ и $\beta=0$, нужно чтобы ξ приняла значение 0, а η приняла значение 2. Так как ξ и η независимы, то

$$P(\alpha=2;\ \beta=0)=P(\xi=0;\ \eta=2)=P(\xi=0)*P(\eta=2)=1/12.$$

Очевидно также $P(\alpha=3;\ \beta=0)=0$.

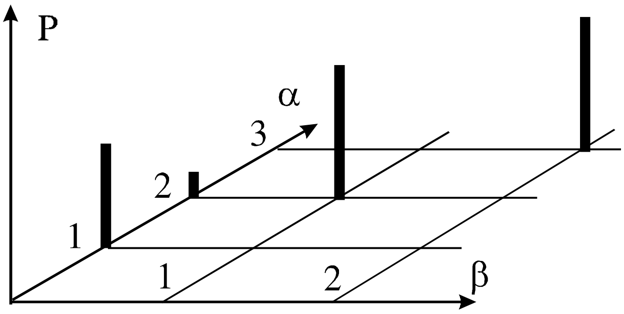

Рис. 2

Построим полигоны условных распределений. Здесь зависимость α от β довольно близка к функциональной: значению $\beta=1$ соответствует единственное $\alpha=2$, значению $\beta=2$ соответствует единственное $\alpha=3$, но при $\beta=0$ мы можем говорить лишь, что α с вероятностью $\frac{3}{4}$ принимает значение 1 и с вероятностью $\frac{1}{4}$ – значение 2.

Пример III.

Рассмотрим закон совместного распределения ξ и η, заданный таблицей

ξ \ η	0	1	2	
1	1/30	3/30	2/30	1/5
2	3/30	9/30	6/30	3/5
3	1/30	3/30	2/30	1/5
	1/6	3/6	2/6	

В этом случае выполняется условие $P(\xi=x_i;\ \eta=y^j)=P(\xi=x_i)*P(\eta=y^j)$, $i=1,2,3\ldots$; $j=1,2,3,\ldots$

Построим законы условных распределений

ξ	1	2	3
$p_{\eta=1}(\xi)=p_{\eta=2}(\xi)=$ $=p_{\eta=3}(\xi)=p_{\eta=4}(\xi)$	1/5	3/5	1/5

Законы условных распределений не отличаются друг от друга при $\eta=1,2,3$ и совпадают с законом распределения случайной величины ξ.

В данном случае ξ и η независимы.

Характеристикой зависимости между случайными величинами ξ и η служит математическое ожидание произведения отклонений ξ и η от их центров распределений (так иногда называют математическое ожидание случайной величины), которое называется коэффициентом ковариации или просто ковариацией.

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

Пусть $\xi=\{x_1,x_2,x_3,\ldots,x_n\}$, $\eta=\{y_1,y_2,y_3,\ldots,y_n\}$. Тогда

$$\mathrm{cov}(\xi;\eta)=\sum_{i=1}^{n}\sum_{j=1}^{k}(x_i-M\xi)(y_j-M\eta)P((\xi=x_i)\cap(\eta=y_j)) \qquad (2)$$

Эту формулу можно интерпретировать так. Если при больших значениях ξ более вероятны большие значения η, а при малых значениях ξ более вероятны малые значения η, то в правой части формулы (2) положительные слагаемые доминируют, и ковариация принимает положительные значения.

Если же более вероятны произведения $(x_i - M\xi)(y_j - M\eta)$, состоящие из сомножителей разного знака, то есть исходы случайного эксперимента, приводящие к большим значениям ξ в основном приводят к малым значениям η и наоборот, то ковариация принимает большие по модулю отрицательные значения.

В первом случае принято говорить о прямой связи: с ростом ξ случайная величина η имеет тенденцию к возрастанию.

Во втором случае говорят об обратной связи: с ростом ξ случайная величина η имеет тенденцию к уменьшению или падению.

Если примерно одинаковый вклад в сумму дают и положительные и отрицательные произведения $(x_i - M\xi)(y_j - M\eta)p_i^j$, то можно сказать, что в сумме они будут “гасить” друг друга и ковариация будет близка к нулю. В этом случае не просматривается зависимость одной случайной величины от другой.

Легко показать, что если

$P((\xi = x_i)\cap(\eta = y_j)) = P(\xi = x_i)P(\eta = y_j)$ $(i = 1,2,\ldots,n;\ j = 1,2,\ldots,k)$,

òî cov(ξ; η)= 0.

Действительно из (2) следует

$$\sum_{i=1}^{n}\sum_{j=1}^{k}(x_i - M\xi)\left(y_j - M\eta\right)P(\xi = x_i)P\left(\eta = y_j\right) =$$

$$= \sum_{i=1}^{n}(x_i - M\xi)P(\xi = x_i)\cdot\sum_{j=1}^{k}\left(y_j - M\eta\right)P\left(\eta = y_j\right) =$$

$$= M(\xi - M\xi)M(\eta - M\eta) = 0\cdot 0 = 0$$

Здесь использовано очень важное свойство математического ожидания: математическое ожидание отклонения случайной величины от ее математического ожидания равно нулю.

Доказательство (для дискретных случайных величин с конечным числом значений).

$$M(\xi - M\xi) = \sum_{i=1}^{n}(x_i - M\xi)P(x_i) = \sum_{i=1}^{n}x_iP(x_i) - M\xi\sum_{i=1}^{n}P(x_i) = M\xi - M\xi = 0$$

Ковариацию удобно представлять в виде

$$\text{cov}(\xi;\ \eta)=M(\xi\eta-\xi M\eta-\eta M\xi+M\xi M\eta)=M(\xi\eta)-M(\xi M\eta)-M(\eta M\xi)+M(M\xi M\eta)=$$
$$=M(\xi\eta)-M\eta M\xi-M\xi M\eta+M\xi M\eta=M(\xi\eta)-M\xi M\eta$$

Ковариация двух случайных величин равна математическому ожиданию их произведения минус произведение математических ожиданий.

Легко доказывается следующее свойство математического ожидания: если ξ и η—независимые случайные величины, то $M(\xi\eta)=M\xi M\eta$. (Доказать самим, используя формулу $M(\xi\eta) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} x_i y_j P(\xi = x_i)P(\eta = y_j)$)

Таким образом, для независимых случайных величин ξ и η $\mathrm{cov}(\xi;\eta)=0$.