## 9. Системы СВ. Совместный, частный и условный законы распределения СВ. Независимость СВ. Свойства ковариации. Коэффициент Корреляции.

### *1.1. Понятие о системе случайных величин*

В практических применениях теории вероятностей приходится иметь дело с задачами, в которых результат опыта описывается не одной случайной величиной, а двумя или более случайными величинами, образующими систему или случайный вектор.

Например, успеваемость наудачу взятого студента характеризуется несколькими оценками, полученными им в ходе экзаменационной сессии;

***Случайным вектором (n-мерной случайной величиной, системой n случайных величин)*** называют упорядоченный набор из n случайных величин ($X_1, X_2, \dots, X_n$).

Одномерные случайные величины $X_1, X_2, \dots, X_n$ называются ***компонентами*** или ***составляющими*** n-мерной случайной величины ($X_1, X_2, \dots, X_n$). Их можно рассматривать как координаты случайной точки или случайного вектора $\overline{X} = (X_1, X_2, \dots, X_n)$ в пространстве n измерений.

Системы случайных величин могут быть ***дискретными, непрерывными*** и ***смешанными*** в зависимости от типа случайных величин, образующих систему.

Функцией распределения системы двух случайных величин $F(x, y)$ называется вероятность совместного выполнения двух неравенств $X < x$ и $Y < y$:

$$F(x, y) = P((X < x)(Y < y)).$$

**Частный (маргинальный) закон распределения** подвектора $\xi_1$ анализируемой многомерной случайной величины $\xi = (\xi_1, \xi_2)$ описывает распределение вероятностей признака $\xi_1$ в ситуации, когда на значения другой части компонент $\xi_2$ не накладывается никаких условий. В дискретном случае соответствующие вероятности определяются по формулам:

$$p_{i\cdot} = P\{\xi_1 = X_i^{(1)0}\} = \sum_j P\{\xi_1 = X_i^{(1)0},\ \xi_2 = X_j^{(2)0}\}; \qquad (5.3)$$

$$p_{\cdot j} = P\{\xi_2 = X_j^{(2)0}\} = \sum_i P\{\xi_1 = X_i^{(1)0},\ \xi_2 = X_j^{(2)0}\}. \qquad (5.3')$$

**5. Условные законы распределения**

*Условным законом распределения одной из случайных величин*, входящей в систему $(X, Y)$, называется ее закон распределения, найденный при условии, что другая случайная величина приняла определенное значение (или попала в какой-то интервал).

Пусть $(X, Y)$ – дискретная двумерная случайная величина и $p_{ij} = P\{X = x_i, Y = y_j\}$, $i = \overline{1, n}$, $j = \overline{1, m}$.

*Условная вероятность того, что случайная величина $Y$ примет значение $y_j$ при условии, что $X = x_i$*, определяется равенством

$$P_{X=x_i}\{Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}}, \ i = \overline{1, n}, \ j = \overline{1, m}. \qquad (12)$$

Совокупность вероятностей (12) представляет собой *условный закон распределения* случайной величины $Y$ при условии $X = x_i$.

Совместным законом распределения дискретных случайных величин ξ и η (или законом их совместного распределения) называется вероятность $P_{ij} = P(\xi = x_i; \eta = y_j)$, определенная на всем множестве упорядоченных пар $(x_i; y_j); i = 1; \dots; k; j = 1; \dots; m$.

Две случайные величины $X$ и $Y$ называются *независимыми*, если независимы все связанные с ними события: например, $\{X < a\}$ и $\{Y < b\}$ или $\{X = xi\}$ и $\{Y = yi\}$ и т.д.

В терминах законов распределения справедливо также следующее определение: две случайные величины $X$ и $Y$ называются *независимыми*, если закон распределения каждой из них не зависит от принятого значения другой.

Ковариация - это количественная мера связи двух случайных величин.

***Корреляционный момент, коэффициент корреляции***

*Корреляционным моментом* (или *ковариацией*) двух случайных величин $X$ и $Y$ называется математическое ожидание произведения отклонений этих СВ от их математических ожиданий и обозначается через $K_{XY}$ или $\operatorname{cov}(X,Y)$.

Таким образом, по определению

$$K_{XY} = \operatorname{cov}(X,Y) = M[(X-m_x)(Y-m_y)]. \tag{25}$$

При этом если $(X,Y)$ – дискретная двумерная СВ, то ковариация вычисляется по формуле

$$K_{XY} = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i-m_x)(y_j-m_y)p_{ij}\,; \tag{26}$$

если $(X,Y)$ – непрерывная двумерная СВ, то

$$K_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-m_x)(y-m_y)f(x,y)dxdy\,. \tag{27}$$

Ковариацию удобно вычислять по формуле

$$K_{XY} = \operatorname{cov}(X,Y) = M(XY) - M(X)\cdot M(Y). \tag{28}$$

Формулу (27) можно записать в виде:

$$K_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy - m_x m_y\,. \tag{29}$$

***Свойства*** ковариации:

1. Ковариация симметрична, т.е. $K_{XY} = K_{YX}$.
2. Дисперсия СВ есть ковариация ее с самой собой, т.е. $K_{XX} = D(X),\ K_{YY} = D(Y)$.
3. Если случайные величины $X$ и $Y$ независимы, то $K_{XY} = 0$.
4. Дисперсия суммы (разности) двух случайных величин равна сумме их дисперсий плюс (минус) удвоенная ковариация этих случайных величин, т.е. $D(X \pm Y) = D(X) + D(Y) \pm 2K_{XY}$.
5. Постоянный множитель можно выносить за знак ковариации, т.е. $K_{cXY} = c\cdot K_{XY} = K_{X,cY}$.
6. Ковариация не изменится, если к одной из СВ (или к обоим сразу) прибавить постоянную, т.е. $K_{X+c,Y} = K_{XY} = K_{X,Y+c} = K_{X+c,Y+c}$.
7. Ковариация двух случайных величин по абсолютной величине не превосходит их средних квадратических отклонений, т.е. $|K_{XY}| \le \sigma_x \cdot \sigma_y$.

Из свойства 3 следует, что если $K_{XY} \ne 0$, то СВ $X$ и $Y$ зависимы. Случайные величины $X$ и $Y$ в этом случае называются *коррелированными*.

*Коэффициентом корреляции* $r_{XY}$ двух СВ $X$ и $Y$ называется отношение их ковариации (корреляционного момента) к произведению их средних квадратических отклонений:

$$r_{XY} = \frac{K_{XY}}{\sigma_x\sigma_y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}\,. \tag{30}$$