

Количественные характеристики случайных переменных

- Математическое ожидание (среднее значение)
- Дисперсия и среднее квадратическое отклонение
- Ковариация и коэффициент корреляции.

Математическое ожидание дискретной случайной переменной

Определение. Математическим ожиданием дискретной случайной переменной называется величина:

$$M(x) = \sum_{i=1}^n P_i x_i \quad (4.1)$$

где: $M(x)$ – математическое ожидание СДП x ,

P_i – вероятность появления в опытах значения x_i ,

x_i – значение дискретной случайной переменной,

n – количество допустимых значений дискретной случайной величины

Математическое ожидание – средневзвешенное значение ДСП, где в качестве веса используется значение вероятности

Дисперсия дискретной случайной переменной

Определение. Дисперсией дискретной случайной переменной называется величина:

$$\sigma^2(x) = \sum_{i=1}^n (x_i - M(x))^2 P(x_i) \quad (4.2)$$

где: $\sigma^2(x)$ – дисперсия случайной переменной x

Дисперсия случайной величины выступает в качестве характеристики разброса возможных ее значений

Положительный корень из дисперсии называют средним квадратическим отклонением или стандартным отклонением, или стандартной ошибкой

Примеры расчета количественных характеристик ДСП

Пример 1. Пусть X_i – результат бросания кубика.

$$A_x = \{1, 2, 3, 4, 5, 6\} \quad P_i = \{1/6, 1/6, 1/6, 1/6, 1/6, 1/6\}$$

Тогда:

$$M(x) = 1/6(1+2+3+4+5+6) = 3.5$$

$$\sigma^2(x) = 1/6[(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2] = 2.92$$

$$\sigma(x) = 1.71$$

Пример 2. Индикатор случайного события

$$I = \begin{cases} 1 & \text{если событие произошло} \\ 0 & \text{если событие не произошло} \end{cases} \quad P_I(t) = \begin{cases} p & \text{если } t = 1 \\ (1-p) & \text{если } t = 0 \end{cases}$$

Математическое ожидание непрерывной случайной переменной

Определение. Математическим ожиданием непрерывной случайной величины X с законом распределения $p_x(t)$ называется величина:

$$M(x) = \int_{-\infty}^{\infty} t p_x(t) dt \quad (4.3)$$

Выражение (4.3) называется первым начальным моментом функции $p_x(t)$

Дисперсия непрерывной случайной переменной

Определение. Дисперсией непрерывной случайной переменной X с функцией плотности вероятности $p_x(t)$ называется выражение:

$$\sigma^2(x) = \int_{-\infty}^{\infty} (t - M(x))^2 p_x(t) dt \quad (4.4)$$

Выражение (4.4) называют вторым центральным моментом функции $p_x(t)$

В общем случае дисперсия случайной переменной определяется как:

$$\sigma^2(x) = M(x - M(X))^2 \quad (4.5)$$

Примеры вычисления

Пример 1. Пусть X НСП с равномерным законом распределения.

$$M(x) = \int_a^b t \frac{1}{(b-a)} dt = \frac{1}{(b-a)} \frac{t^2}{2} = \frac{(b+a)}{2}$$

$$\sigma^2(x) = \int_a^b \left(t - \frac{(a+b)}{2} \right)^2 \frac{1}{(b-a)} dt = \frac{1}{3(b-a)} \left(t - \frac{(a+b)}{2} \right)^3 = \frac{(b-a)^2}{12}$$

Самостоятельно вычислить математическое ожидание и дисперсию НСП с нормальным законом распределения

Понятие ковариации двух случайных переменных

По определению ковариацией двух случайных переменных X и Y есть:

$$\text{COV}(x,y)=M((x-M(x))(y-M(y))) \quad (4.6)$$

Значение ковариации отражает наличие связи между двумя случайными переменными

Если $\text{COV}(x,y)>0$, связь между X и Y положительная

Если $\text{COV}(x,y)<0$, связь между X и Y отрицательная

Если $\text{COV}(x,y)=0$, X и Y независимые переменные

Область возможных значений ковариации – вся числовая ось

Понятие коэффициента корреляции двух случайных переменных

Недостатки ковариации в том, что ее значения зависят от масштаба измерения переменных и наличия размерности

Недостатки устраняется путем деления значения ковариации на значения стандартных отклонений переменных:

$$\rho = \frac{COV(X, Y)}{\sigma(X)\sigma(Y)} \quad (4.7)$$

Выражение (4.7) называют коэффициентом корреляции двух случайных переменных

Коэффициент корреляции изменяется в пределах $[-1; 1]$ и является безразмерной величиной

Основные свойства количественных характеристик

■ Свойства математического ожидания.

$$M(c) = c$$

$$M(c_1x_1 + c_2x_2) = c_1x_1 + c_2x_2$$

Пример. $M(Y) = M(f(\vec{X}) + \varepsilon) = M(f(\vec{X})) + M(\varepsilon) = M(f(\vec{X}))$

Свойства дисперсий.

$$\sigma^2(c) = 0$$

$$\sigma^2(c + x) = \sigma^2(x)$$

$$\sigma^2(c_1x_1 + c_2x_2) = c_1\sigma^2(x_1) + c_2\sigma^2(x_2) + 2c_1c_2\text{Cov}(x_1, x_2)$$

В общем случае:

$$\sigma^2(\sum c_i x_i) = c^T \text{Cov}(XX) c$$

Основные свойства количественных характеристик

■ Свойства ковариаций.

$$\text{Cov}(x, y) = \text{Cov}(y, x)$$

$$\text{Cov}(c_1 x_1 + c_2 x_2) = c_1 c_2 \text{Cov}(x_1, x_2)$$

$$\text{Cov}(cx) = 0$$

$$\text{Cov}(x+c, y) = \text{Cov}(x, y)$$

$$\text{Cov}(x+y, z) = \text{Cov}(x, z) + \text{Cov}(y, z)$$

$$\text{Cov}(x, x) = \sigma^2(x)$$

Доказательства этих свойств проведите самостоятельно!

Связь между случайными переменными

Случайный вектор и его количественные характеристики.

Пусть опыт – инвестирование средств на некоторый период времени в рисковые активы $A=\{a_1, a_2, \dots, a_n\}$.

Рисковый характер актива означает, что значения доходности на них являются случайными величинами $r(a_1), r(a_2), \dots, r(a_n)$.

Определение. Вектор, компонентами которого являются случайные величины, называется случайным вектором.

Пример 1. Вектор доходностей по рисковым активам

$$R=\{r(a_1), r(a_2), \dots, r(a_n)\}^T. \quad (4.1)$$

Пример 2. Опыт – бросание игральной кости.

Пусть X – количество очков на верхней грани кости, а Y – количество очков на его нижней грани. Тогда вектор $Z=\{X, Y\}^T$ – пример случайного вектора.

Связь между случайными переменными

Случайный вектор и его количественные характеристики.

Пусть $m_i = M(r(a_i))$ – ожидаемое значение доходности актива a_i ,

$\sigma_i^2 = M(r(a_i) - m_i)^2$ – дисперсия доходности актива a_i ,

$\sigma_{ij} = \text{Cov}(r(a_i), r(a_j))$ – ковариация между активами a_i, a_j .

Тогда вектор

$$M = \{m_1, m_2, \dots, m_n\}^T = M(R) \rightarrow (4.2)$$

является первой основной характеристикой случайного вектора (4.1).

Замечание. Вектор M является константой.

Ковариационная матрица вида:

$$\sigma_{RR} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{pmatrix}$$

Является второй основной характеристикой случайного вектора \vec{R}

Связь между случайными переменными

Параметрическая модель Марковца фондового рынка.

По предложению Марковца компоненты вектора R рассматривается как характеристики привлекательности каждого рискованного актива, а диагональные элементы ковариационной матрицы – как характеристики риска инвестирования в эти активы.

Параметрической моделью Марковца называется следующая тройка:

$$\{A, M, \sigma_{rr}\} \quad (4.4)$$

Для формирования индивидуального пакета акций из списка A ничего больше не требуется.

Эта модель является инструментом брокерской деятельности.

Выборка и ее свойства

- Задачи математической статистики.
 1. Оценивание (приближенное определение) параметров законов распределения и самих законов.
 2. Проверка различных гипотез относительно законов распределения или значений их параметров.

Далее будем рассматривать случайные величины с законом распределения $R(t, a_1, a_2, \dots, a_n)$, где $A = \{a_1, a_2, \dots, a_n\}^T$ вектор столбец параметров распределения.

Выборка и ее свойства

Определение. Выборка – это случайный вектор, составленный из результатов наблюдений, каждое из которых суть случайная величина.

$$Y = \{y_1, y_2, \dots, y_n\}$$

$$y_1 = t_1; \quad P_y(t_1, a_1, a_2, \dots, a_k);$$

$$y_2 = t_2; \quad P_y(t_2, a_1, a_2, \dots, a_k);$$

.....

$$y_n = t_n; \quad P_y(t_n, a_1, a_2, \dots, a_k);$$

Выборка и ее свойства

■ Свойства случайной выборки.

1. Каждый элемент выборки есть случайная величина с тем же законом распределения, что и случайная величина Y .
2. Все значения, входящие в выборку независимые величины.

Тогда для них справедлива теорема умножения вероятностей:

$$P_y(y_1, y_2, \dots, y_n | A) = P_y(t_1, A) P_y(t_2, A) \dots P_y(t_n, A)$$

Это выражение – закон распределения выборки.

Задача заключается в том, чтобы найти процедуры, с помощью которых можно найти значения параметров распределения.

$$A = F(y_1, y_2, \dots, y_n)$$

Свойства оценок параметров распределения.

1. Оценка представляет собой частный случай случайной величины.

Например. Рассмотрим оценку математического ожидания в виде среднего значения:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Любую случайную величину можно представить в виде:

$$X_i = \mu + U_i$$

где: U_i – случайная величина,

μ – константа равная математическому ожиданию X_i .

$$X = 1/n \sum (\mu + U_i) = \mu + U$$

Свойства оценок параметров распределения.

1. Несмещенность оценки.

$$M(\tilde{a}) = a$$

Процедуры, которые дают такие оценки будем называть несмещенными.

Замечание. Несмещенных процедур может быть много.

Пример. Оценка среднего значения. $\bar{X} = 1/n \sum x_i$

Эта процедура несмещенная т.к.

$$M(X) = M(\mu + U) = M(\mu) + M(U) = \mu$$

Вопрос. Можно ли найти иную несмещенную процедуру?

Пусть имеем выборку из двух значений x_1 и x_2 , следовательно:

$$M(x_1) = M(x_2) = \mu \text{ и } \sigma(x_1) = \sigma(x_2) = \sigma$$

Пусть такой процедурой будет: $Z = \lambda_1 x_1 + \lambda_2 x_2$

$$M(Z) = M(\lambda_1 x_1 + \lambda_2 x_2) = (\lambda_1 + \lambda_2) \mu$$

Вывод. Все процедуры, для которых $\lambda_1 + \lambda_2 = 1$ дают несмещенные оценки среднего значения.

Свойства оценок параметров распределения.

2. Эффективность оценки.

Определение. Оценка называется эффективной среди всех оценок параметра, если она имеет минимальную дисперсию среди всех возможных оценок: $\sigma^2(\tilde{a}) = \min$.

Задача. При каких значениях λ_1 и λ_2 оценка среднего значения будет эффективной?

Найдем минимум дисперсии Z .

$$\sigma^2(Z) = \sigma^2(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1^2 \sigma^2(x_1) + \lambda_2^2 \sigma^2(x_2) = (\lambda_1^2 + \lambda_2^2) \sigma^2$$

Учитывая, что $(\lambda_1 + \lambda_2) = 1$ или $\lambda_2 = (1 - \lambda_1)$, получим:

$$\sigma^2(Z) = (\lambda_1^2 + (1 - \lambda_1)^2) \sigma^2$$

Тогда: $\partial \sigma^2 / \partial \lambda_1 = (2\lambda_1 - 2(1 - \lambda_1)) \sigma^2$

Откуда $\lambda_1 = 1/2$.

Вторая производная положительна, следовательно, это минимум. Аналогичным образом можно показать, что известная оценка дисперсии также не смещена и эффективна.