Основные понятия экономической статистики

План:

1. Понятие случайной величины
2. Закон распределения ДСВ
3. Числовые характеристики ДСВ
4. Генеральная и выборочная совокупность
5. Ковариация, корреляция и дисперсия

Случайная величина

Случайной называют величину, которая в результате испытания примет одно и только одно возможное значение, заранее не известное и зависящее от случайных причин

Примеры

1. Число родившихся мальчиков среди ста новорожденных есть случайная величина, которая имеет следующие возможные значения: 0, 1, 2,...,100.

2. Расстояние, которое пролетит снаряд при выстреле из орудия, есть случайная величина. Действительно, расстояние зависит не только от установки прицела, но и от многих других причин (силы и направления ветра, температуры и т. д.), которые не могут быть полностью учтены. Возможные значения этой величины принадлежат некоторому промежутку (а, Ь).

Дискретные и непрерывные случайные величины

- **Дискретной** называют случайную величину, которая принимает отдельные, изолированные значения с определенными вероятностями
- **Непрерывной** называют случайную величину, которая может принимать все значения из некоторого конечного или бесконечного промежутка

Закон распределения вероятностей ДСВ

Законом распределения дискретной случайной величины (ДСВ) называют соответствие между возможными значениями и их вероятностями (сумма вероятностей равна единице)

X	1	3	4	7
P	0,2	0,1	0,3	0,4

Пример

В денежной лотерее выпущено 100 билетов. Разыгрывается один выигрыш в 50 млн.сум и десять выигрышей по 1 млн. сум. Найти закон распределения случайной величины **X** — стоимости возможного выигрыша для владельца одного лотерейного билета.

Решение

X	50	1	0
P	0,01	0,1	0,89

P = 0,01+0,1+0,89 = 1.

Числовые характеристики ДСВ

Математическим ожиданием ДСВ называют сумму произведений всех ее возможных значений на их вероятности:

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

$$M(X) = x_1\, p_1 + x_2\, p_2 + \ldots + x_n\, p_n$$

Пример

Найти математическое ожидание случайной величины Х:

X	3	5	2
P	0,1	0,6	0,3

$M(X) = 3*0,1 + 5*0,6 + 2*0,3 = 3,9.$

Задание

Найти математическое ожидание случайной величины X:

X	2	5	8
P	0,2	0,5	0,3

$M(X) = 2*0{,}2 + 5*0{,}5 + 8*0{,}3 = 5{,}3.$

Вероятностный смысл

Вероятностный смысл математического ожидания таков: математическое ожидание приближенно равно (тем точнее, чем больше число испытаний) **среднему арифметическому** наблюдаемых значений случайной величины

Свойства математического ожидания

1. Математическое ожидание постоянной величины равно самой постоянной:

$$M(C) = C$$

2. Постоянный множитель можно выносить за знак математического ожидания:

$$M(CX) = C*M(X)$$

Свойства математического ожидания

3. Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:

$$M(X*Y) = M(X)*M(Y)$$

4. Математическое ожидание суммы двух случайных величин равно сумме их математических ожиданий:

$$M(X+Y) = M(X) + M(Y)$$

Дисперсия ДСВ

Можно указать такие случайные величины, которые имеют одинаковые математические ожидания, но различные возможные значения

X	-0,01	0,01
P	0,5	0,5

M(X) = -0,01*0,5 + 0,01*0,5 = 0.

Y	-100	100
P	0,5	0,5

M(Y) = -100*0,5 + 100*0,5 = 0.


Black 10k
Blue 10k

Дисперсия ДСВ

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

$$D(X) = M[X - M(X)]^2$$

Дисперсия

Дисперсия это **оценка рассеяния** возможных значений случайной величины вокруг ее среднего значения.

Формула для вычисления дисперсии:

$$D(X) = M(X^2) - [M(X)]^2$$

Пример

Найти дисперсию случайной величины X, которая задана следующим законом распределения:

X	2	3	5
P	0,1	0,6	0,3

M(X) = 2*0,1 + 3*0,6 + 5*0,3 = 3,5

X^2	4	9	25
P	0,1	0,6	0,3

$M(X^2)$ = 4*0,1 + 9*0,6 + 25*0,3 = 13,3

$D(X) = M(X^2) – [M(X)]^2 = 13{,}3 - (3{,}5)^2 = 1{,}05$

Свойства дисперсии

1. Дисперсия постоянной величины равна нулю:

$$D(C) = 0$$

2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат:

$$D(C*X) = C^2*D(X)$$

Свойства дисперсии

3. Дисперсия суммы двух независимых случайных величин равна сумме дисперсий этих величин:

$$D(X + Y) = D(X) + D(Y)$$

4. Дисперсия разности двух независимых случайных величин равна сумме дисперсий этих величин:

$$D(X - Y) = D(X) + D(Y)$$

Среднее квадратическое отклонение

Средним квадратическим отклонением случайной величины X называют квадратный корень из дисперсии:

$$\sigma(X) = \sqrt{D(X)}$$

Пример

Дисперсия случайной величины равна

D(X) = 1,05

Найти среднее квадратическое отклонение

$$\sigma(X) = \sqrt{D(X)} = \sqrt{1,05} = 1,02$$

Среднее квадратическое отклонение

Дисперсия имеет размерность, равную квадрату размерности случайной величины. Так как среднее квадратическое отклонение (СКО) равно квадратному корню из дисперсии, то размерность СКО совпадает с размерностью X. Поэтому в тех случаях, когда желательно, чтобы оценка рассеяния имела размерность случайной величины, вычисляют среднее квадратическое отклонение, а не дисперсию

Задание

Найти математическое ожидание, дисперсию и среднее квадратическое отклонение дискретной случайной величины, зная закон ее распределения:

X	6+к	3+к	1+к
P	0,2	0,3	0,5

Терминология

- Генеральная совокупность (population) это совокупность всех объектов
- Выборочная совокупность или просто выборка (sample) это совокупность случайно отобранных из генеральной совокупности объектов
- Объемом совокупности называют число объектов этой совокупности (объем генеральной совокупности - population size, объем выборки - sample size)

Выборочный метод

- **Повторной** называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность
- **Бесповторной** называют выборку, при которой отобранный объект в генеральную совокупность не возвращается
- Выборка должна правильно представлять пропорции генеральной совокупности, т.е. она должна быть **представительной (репрезентативной)**
- В силу закона больших чисел можно утверждать, что выборка будет репрезентативной, если ее осуществить случайно: каждый объект выборки отобран случайно из генеральной совокупности, если все объекты имеют одинаковую вероятность попасть в выборку

Функциональная, статистическая и корреляционная зависимости

- Две случайные величины могут быть связаны либо **функциональной** зависимостью, либо зависимостью другого рода, называемой **статистической**, либо быть **независимыми**.
- Функциональная зависимость: Если каждому значению случайной величины **X** соответствует одно возможное значение случайной величины **Y**, то **Y** называют функцией случайного аргумента **X**:

$$\mathbf{Y = \varphi(X)}$$

Примеры (функциональная зависимость)

1. $y = ax + b$
2. $y = x^2$
3. $y = \sin(x)$

Статистическая зависимость

- Статистической называют зависимость, при которой изменение одной из величин влечет изменение **распределения** другой
- В частности, статистическая зависимость проявляется в том, что при изменении одной из величин изменяется **среднее** значение другой; в этом случае статистическую зависимость называют **корреляционной**

Пример

Зависимость между урожайностью хлопка и количеством вносимых удобрений является статистической. Действительно, увеличение количества вносимых удобрений приводит к увеличению урожайности хлопка **в среднем**

Корреляционный анализ

Пусть изучается зависимость между факторами **X** и **Y**. В результате **n** независимых опытов получены **n** пар чисел:

$$(x_1; y_1), (x_2; y_2), \ldots, (x_n; y_n)$$

X	x_1	x_2	…	x_n
Y	y_1	y_2	…	y_n

Коэффициент ковариации

Является мерой взаимосвязи между факторами:

$$\operatorname{cov}(x, y) = \frac{1}{n}\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$$

Пример 1

Найти коэффициент ковариации:

Количество внесенных удобрений	150	160	180	200	250
Урожайность	25	24	26	27	28

$$\bar{x} = (150 + 160 + 180 + 200 + 250)/5 = 188$$

$$\bar{y} = (25 + 24 + 26 + 27 + 28)/5 = 26$$

$$\operatorname{cov}(x,y) = ((150-188)(25-26) + (160-188)(24-26) + (180-188)(26-26) + (200-188)(27-26) + (250-188)(28-26))/5 = 46$$

Тесноту связи между двумя взаимозависимыми рядами характеризует коэффициент линейной корреляции, который показывает, существует ли и насколько велика связь изучаемых явлений

$$r = \frac{\sum (x-\overline{x})(y-\overline{y})}{\sqrt{\sum (x-\overline{x})^2 \sum (y-\overline{y})^2}}$$

Пример 2

Найти коэффициент корреляции:

Количество внесенных удобрений	150	160	180	200	250
Урожайность	25	24	26	27	28

	Количество внесенных удобрений	Урожайность	x-xcp	y-ycp	(x-xcp)(y-ycp)	(x-xcp)^2	(y-ycp)^2	r	D
1	150	25	-38	-1	38	1444	1		
2	160	24	-28	-2	56	784	4		
3	180	26	-8	0	0	64	0		
4	200	27	12	1	12	144	1		
5	250	28	62	2	124	3844	4		
Сумма	940	130			**230**	**6280**	**10**		
Среднее	**188**	**26**			**46**		250,599	**0,92**	**0,84**

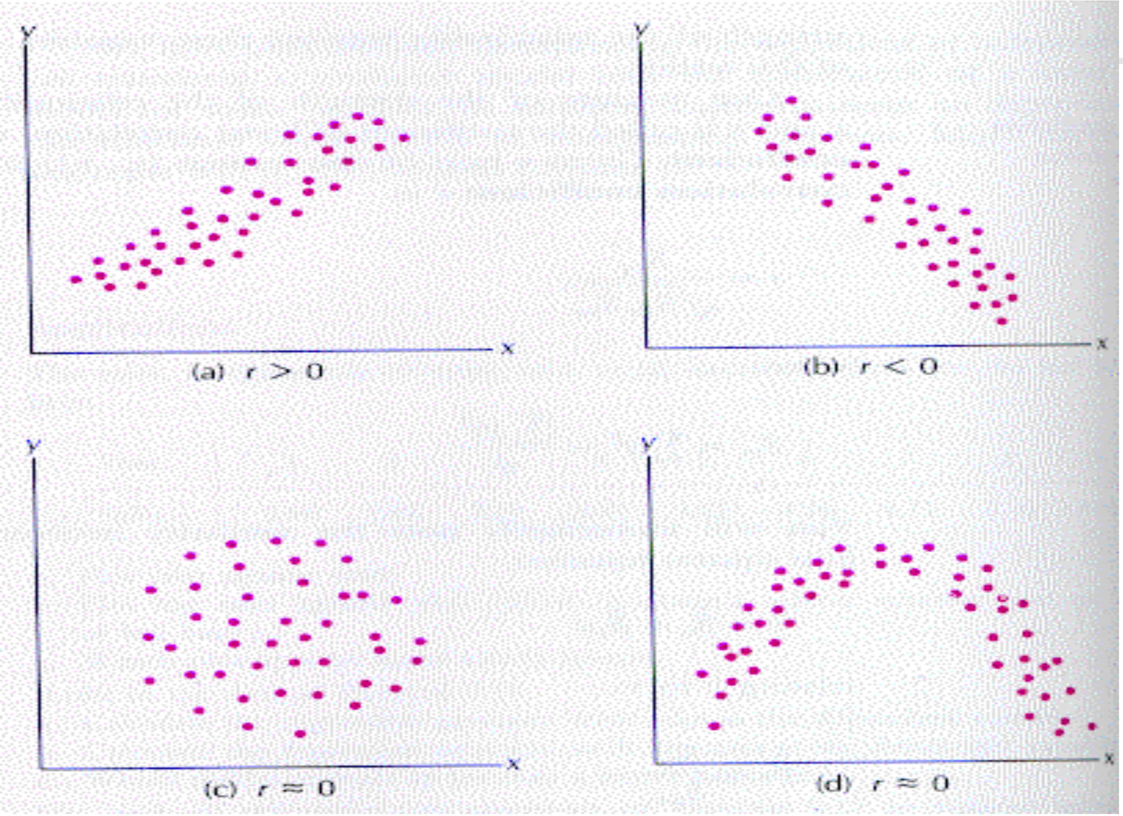
Коэффициент корреляции

- Коэффициент корреляции r может принимать значения от -1 до 1. Если $r<0$, то связь обратная, если $r>0$, то связь прямая

Коэффициент корреляции

Чем ближе значение коэффициента корреляции к **-1** или **+1**, тем теснее связь между факторами, и наоборот, чем ближе значение коэффициента корреляции к **0**, тем слабее связь между факторами

Свойства r



Коэффициент корреляции

- Коэффициент корреляции указывает следующую степень связи:
- 0 ÷ ±0,15 – отсутствие связи
- ±0,16 ÷ ±0,20 – плохая
- ±0,21 ÷ ±0,30 – слабая
- ±0,31 ÷ ±0,40 – умеренная
- ±0,41 ÷ ±0,60 – средняя
- ±0,61 ÷ ±0,80 – высокая
- ±0,81 ÷ ±0,9 – очень высокая
- ±0,91 ÷ ±1,0 – полная

Коэффициент детерминации

При анализе взаимосвязи между факторами также вычисляют коэффициент детерминации (r-квадрат)

$$D = r^2$$

Коэффициент детерминации

Коэффициент детерминации показывает, какое влияние оказывают выбранные факторы на результативный показатель

Пример 3

Например, если **r = 0,92** (**r** - коэффициент корреляции), то **D=0,84**, то есть величина результативного показателя на **84% (0,84·100)** зависит от изменения исследуемых факторов и на **16%** от остальных факторов

СПАСИБО ЗА ВНИМАНИЕ!

☎ + 998 71 237 1948

✉ s.mirzaev@tiiame.uz