**16. Корреляция, двумерное нормальное распределение**

Рассмотрим некоторые понятия, необходимые для изучения взаимной зависимости между случайными величинами. Совместное распределение двух случайных величин в непрерывном случае мы определили через плотность совместного распределения. Но можно построить также и функцию совместного распределения, как для непрерывных, так и дискретных случайных величин.

**Определение.** Функцией совместного распределения случайных величин называется

Для дискретных случайных величин функция распределения очевидным образом выражается через вероятности (см. Параграф 11). Для непрерывных случайных величин задана плотность совместного распределения , через которую можно записать

Отметим некоторые свойства совместной функции распределения, прямо следующие из свойств вероятности:

при всех *x,y*;

;

имеет также место свойство, согласно которому совместная функция распределения зависит от *x,y* только через функции распределения, :

),

где – функция распределения, заданная на единичном квадрате; в частности, для независимых случайных величин

то есть здесь .

Для измерения взаимной зависимости между случайными величинами часто бывает полезен следующий показатель.

**Определение.** Ковариацией случайных величин называется

**Теорема 16.1.** Ковариация обладает следующими свойствами:

1. ;
2. ;
3. для независимых случайных величин

**Доказательство.** Первые четыре свойства вполне очевидно следуют из определения, равенство 3) уже встречалось в Параграфе 11 (следствие из Теоремы 11.2.). Чтобы доказать пятое свойство, рассмотрим квадратичное по *x* (*x* - произвольное вещественное число) выражение

поскольку оно всегда неотрицательно, квадратное уравнение

не имеет вещественных корней, значит, его дискриминант отрицателен,

следовательно,

откуда и следует доказываемое неравенство.

Это неравенство подсказывает идею ввести нормированную ковариацию: если разделить на произведение среднеквадратичных отклонений, максимальное значение такого коэффициента станет равным 1. Такая нормированная ковариация называется корреляцией.

**Определение.** Коэффициентом корреляции случайных величин называется

Свойства коэффициента корреляции:

2. для независимых случайных величин
3. максимальное значение коэффициента корреляции, , достигается для линейно связанных случайных величин, .

Докажем последнее свойство: подставляя в формулу для , находим

Очень часто возникает необходимость перейти от случайных величин к новым, полученным в результате некоторого функционального преобразования. Рассмотрим, по каким правилам при этом изменяется совместное распределение. Обозначим для удобства пару случайных величин через имеющих плотность совместного распределения , и пусть – две заданные функции, с помощью которых определим новые случайные величины:

.

Функция совместного распределения для них:

сделав замену переменных,

преобразуем интеграл согласно правилам замены переменных в двойном интеграле:

где - обратное по отношению к ) отображение, а - его якобиан,

Дифференцируя по полученное выражение для функции распределения, находим плотность совместного распределения,

**Определение.** Двумерным нормальным распределением называется совместное распределение двух случайных величин, имеющее плотность

здесь - заданные вещественные числа, положительные и

**Теорема 16.2** Если случайные величины имеют двумерное нормальное распределение, то каждая в отдельности имеет нормальное распределение, причем и - коэффициент корреляции.

**Доказательство.** Согласно основному свойству совместной плотности (Параграф 11),

Подставим сюда плотность двумерного гауссовского распределения ( ), в которой очевидным образом сгруппируем слагаемые и выделим полный квадрат по переменной *x:*

и тогда преобразуем интеграл:

Интеграл справа представляет собой интеграл по всей вещественной оси от некоторой гауссовской плотности, а потому он равен 1 (условие нормировки). Оставшееся выражение и есть плотность распределения Точно также выводится распределение для , таким образом, первая часть Теоремы, касающаяся математических ожиданий и дисперсий, доказана.

Вычислим коэффициент корреляции,

здесь использовано полученное выше представление совместной плотности. Интеграл по переменной *x* преобразуем следующим образом:

,

тогда

и подставляя в ( ), убеждаемся, что действительно

Отметим одно свойство нормального распределения (как будет видно в будущем -очень важное свойство): если имеют двумерное нормальное распределение и они некоррелированы, то независимы. Это прямо следует из формулы ( ), поскольку при *r* = 0 совместная плотность превращается в произведение двух плотностей:

много интересного материала по двумерному распределению см Крамер гл 21

нарисовать картинку с двумерной плотностью в аксонометрии