

Понятие о корреляционном и регрессионном анализах

Выполнила Потеха Ольга
Группа ОДЛ 15 с
2016 год

регрессионной связи

- Две (или более) случайных величины
- 1. Могут быть связаны *функциональной* зависимостью – для каждой независимой переменной X существует вполне определенное значение зависимой переменной Y . Строгая функциональная зависимость реализуется на практике редко, т.к. обе величины подвергаются еще и влиянию случайных факторов.
- 2. Могут быть связаны *статистической* зависимостью – изменение одной случайной величины приводит к изменению распределения другой случайной величины. Если статистическая зависимость проявляется в том, что при изменении одной величины изменяется *среднее значение* другой, такую зависимость называют *корреляционной*.
- 3. Могут быть независимы.

- При изучении конкретных зависимостей между случайными величинами вводят понятия:
- факторные признаки или факторы - независимые или объясняющие переменные, причины. Могут быть случайными и неслучайными. Часто обозначаются X ;
- результативные признаки или показатели - объясняемые или зависимые переменные. Являются случайными. Часто обозначаются Y .
- Иногда X и Y можно менять местами (т.е. не только изменение X вызывает изменение Y , но и наоборот, изменение Y вызывает изменение X).
- Функциональная и корреляционная зависимость отличаются тем, что при функциональной зависимости, зная X , можно вычислить величину Y . При корреляционной зависимости устанавливается лишь тенденция изменения Y при изменении X .
- Корреляционный и регрессионный анализы имеют общие методы обработки данных, но отличаются своими целями. В корреляционном анализе оценивается наличие и глубина (сила) статистической связи, в регрессионном анализе оценивается форма статистической связи между случайными величинами.
- Если не известно, какой из признаков зависимый, а какой - независимый, или же это безразлично, то X и Y равноправны, т.е. каждый из признаков может рассматриваться как независимый или как зависимый. В этом случае говорят,

- Регрессия - это односторонняя стохастическая зависимость, когда одна из переменных служит причиной для изменения другой.
- Например, при изучении потребления электроэнергии (Y) в зависимости от объема производства (X) речь идет об односторонней связи, следовательно, о регрессии.
- Существуют особенности, связанные с постановкой задачи:
 - · если изучают стохастическую зависимость Y от X , то устанавливают регрессию Y на X , т.е. $Y=f(X)$;
 - · если изучают стохастическую зависимость X от Y - то устанавливают регрессию X на Y , т.е. $X=g(Y)$.
- Например, изучается влияние стоимости товара на спрос и влияние спроса на стоимость товара. Здесь и стоимость, и спрос могут быть зависимой и независимой переменными в зависимости от постановки задачи.
- Могут быть ситуации, когда обратная регрессия не имеет физического смысла, например, урожайность зависит от количества осадков, обратная зависимость бессмысленна.

корреляций

- При изучении взаимосвязи факторных и результативных признаков могут быть следующие случаи:
 - а) X и Y — случайные величины;
 - б) X - неслучайная величина, Y - случайная величина.
- **Виды корреляции классифицируются по следующим признакам:**
 - а) по характеру корреляции:
 - · положительная (или равнонаправленная, прямая корреляция);
 - · отрицательная (или обратная корреляция);
 - б) по числу переменных:
 - · простая или парная корреляция (две переменных X и Y);
 - · множественная корреляция (рассматривается связь более двух переменных);
 - · частная корреляция (рассматривается связь между двумя переменными при фиксированном влиянии других переменных);
 - в) по форме связи:
 - · линейная корреляция;
 - · нелинейная корреляция;
 - г) по типу связи признаков:
 - · непосредственная корреляция;
 - · косвенная корреляция;
 - · ложная корреляция.

классифицируются по следующим признакам:

- а) по числу переменных, учитываемых в регрессии:
 - · простая регрессия (парная – рассматриваются две переменных);
 - · множественная, или частная регрессия (рассматривается более двух переменных);
- б) по форме зависимости между факторными и результирующими признаками:
 - · линейная регрессия (признаки связаны линейной зависимостью);
 - · нелинейная регрессия (признаки связаны нелинейной зависимостью);
- в) по характеру регрессии (имеет смысл только для простой линейной регрессии):
 - · положительная регрессия;
 - · отрицательная регрессия;
- г) по типу связи факторных и результирующих признаков:
 - · непосредственная регрессия - причина прямо воздействует на следствие;
 - · косвенная регрессия, Y и X не состоят в прямой зависимости, а определяются общей для них причиной через третью переменную;
 - · нонсенс-регрессия (абсурдная).

регрессионного анализа

- 1. Задачи корреляционного анализа:
 - а) определяет степень связи двух и более признаков;
 - б) определяет факторы оказывающее наибольшее влияние на результирующий признак Y .
- 2. Задачи регрессионного анализа:
 - а) устанавливает форму зависимости (для случая парной регрессии – убывающая или возрастающая);
 - б) определяет вид функции регрессии;
 - в) оценивает неизвестные значения зависимой переменной Y (можно воспроизвести значение Y при заданных значениях X внутри рассматриваемого интервала (интерполяция) и вне интервала (экстраполяция)).
- Ход рассуждений, постановка задачи, получаемые результаты в корреляционном и регрессионном анализе различны, но очень часто эти два вида анализа проводятся параллельно на одном и том же массиве исходных данных.

Корреляция

- Корреляционный анализ используется для численной оценки силы связи между случайными величинами (признаками), которые характеризует некоторый реальный процесс.
- В общем виде задача выявления и оценки силы стохастической связи не решена до сих пор. Корреляционная связь это частный случай стохастической зависимости, которая существует между значениями одного из признаков (принятого за независимый) и групповыми средними значениями другого (зависимого) признака.
- Чаще всего корреляционная связь характеризуется выборочным коэффициентом корреляции r , который характеризует степень линейной функциональной зависимости между случайными величинами X и Y .