

АНАЛИЗ ДИСПЕРСИОННЫЙ (ANOVA) - статистический метод, предназначенный для исследования причинной связи между переменной зависимой (см.) и одной или несколькими независимыми переменными (факторами). По числу независимых переменных, различают однофакторный А.Д., двухфакторный А.Д. и т.д. (не путать с Анализом факторным!) Зависимая переменная может быть только количественной, в то время как независимые переменные могут быть представлены номинальными, порядковыми или количественными шкалами (последние должны быть дискретными или сгруппированными в интервалы).

Первоначально А.Д. был разработан для обработки данных, полученных в ходе специально поставленных экспериментов (см.), и считался единственным методом, корректно исследующим причинные связи. Однако в последние годы, вследствие общей либерализации взглядов на природу статистического анализа, он применяется также к данным, собранным в результате выборочных обследований, если численность выделяемых групп не слишком велика.

В основе метода лежит сравнение средних значений зависимой переменной для групп, образованных комбинациями факторов (сами значения факторов не рассматриваются).

Это обстоятельство позволяет некоторым авторам рассматривать А.Д. как обобщение t -критерия Стьюдента, предназначенного для сравнения средних значений переменной в двух группах.

В А.Д. все различия в значениях зависимой переменной (y) объясняются двумя причинами: ее "собственной" или "естественной" изменчивостью, а также изменчивостью, вызванной влиянием независимых переменных, которые в данном случае называются факторами (см. Дисперсия объясненная). Соответственно сумма квадратов (см.) зависимой переменной $SS_{\text{общая}} = \sum (y_{ij} - \bar{y})^2$, где j - номер группы, к которой принадлежит объект с номером i , может быть разложена на две составляющие, одна из которых отражает собственную изменчивость y , а вторая - изменчивость, вызванную влиянием факторов:

$$SS_{\text{общая}} = SS_{\text{внутригрупповая}} + SS_{\text{межгрупповая}} .$$

Внутригрупповая сумма квадратов

$$SS_{\text{внутригрупповая}} = \sum (y_{ij} - \bar{y}_j)^2 ,$$

где \bar{y}_j - среднее арифметическое (см.) переменной y для группы j , в которую входит объект с номером i , рассчитывается внутри групп, образованных факторами, и измеряет степень разброса зависимой переменной в однородных условиях. Межгрупповая сумма квадратов $SS_{\text{межгрупповая}}$ рассчитывается для средних значений y в разных группах и

отражает различия, вызванные влиянием факторов.

Модель однофакторного А.Д. предполагает, что среднее значение зависимой переменной y в группе с номером j (y_j) зависит, во-первых, от среднего значения y по всей совокупности i , во-вторых, от эффекта j -го значения фактора x_j , который обозначается α_j : $y_j = \mu + \alpha_j$. Если фактор x не влияет на зависимую переменную y , то все $\alpha_j = 0$ и, следовательно, все $y_j = \mu$, т.е. не отличаются друг от друга. Проверка гипотезы о влиянии фактора на зависимую переменную сводится к проверке нулевой гипотезы (см.) о том, что все α_j равны нулю

(H_0 : $\alpha_j = 0$ для всех значений j), означающей также, что все групповые средние равны друг другу, против альтернативной гипотезы, состоящей в том, что хотя бы для некоторых значений фактора x_j эффекты α_j отличны от нуля (H_1 : $\alpha_j \neq 0$ хотя бы для некоторых j). Принятие нулевой гипотезы означает признание независимости переменной y от фактора x ; ее отклонение (и, соответственно, принятие альтернативной гипотезы) может служить подтверждением исследуемой причинной зависимости.

Результаты однофакторного А.Д. представляются в виде стандартной таблицы, известной как таблица А.Д. (См. Таблицу 1 Приложения.)

Если H_0 верна, то F -отношение имеет распределение Фишера с $df_{\text{мг}} = k - 1$ и $df_{\text{вн}} = n - k$ степенями свободы (см.). Нулевая гипотеза должна быть отклонена, если вычисленное значение F превысит критическое значение $F_{1-\alpha}$, где α - заданный уровень значимости (см.).

Многофакторный А.Д. концептуально не отличается от однофакторного.

В двух- и многофакторных моделях проверяется ряд гипотез о влиянии на зависимую переменную каждого отдельно взятого фактора, а также их всевозможных сочетаний, что математически выражается в гипотезах о равенстве нулю прямых эффектов факторов и эффектов их взаимодействия. Для проверки таких гипотез межгрупповая сумма квадратов ($SS_{\text{межгрупповая}}$), представляющая вариацию зависимой переменной под общим влиянием всех факторов, делится на более мелкие составляющие, каждая из которых представляет прямой эффект одного из факторов или один из эффектов взаимодействия.

А.Д. не позволяет судить о том, как именно влияют факторы на значение зависимой переменной, т.е. в каких конкретно группах средние достоверно различаются. Для проверки конкретных гипотез о характере таких различий используются метод парных сравнений (см.) и методы множественных сравнений (см.).

АНАЛИЗ ДИСПЕРСИОННЫЙ

Автор: словарь
01.09.2007 06:42 -

О.В. Терещенко