

АНАЛИЗ СОБЫТИЙНЫЙ - анализ данных об истории событий, т.е. о времени и обстоятельствах (ковариатах) их наступления в границах наблюдаемого периода времени и о самих этих событиях. Под событием понимается изменение статуса, измеряемого дискретной переменной, принимающей счетное количество взаимно исключающих значений. Например, в исследовании брачного поведения событиями могут быть вступление в брак, развод, повторный брак, вдовство и т.д. Единицами А.С. могут выступать индивиды, группы и организации, страны, политические объединения, культурные процессы.

Для предсказания времени наступления события строят модель, подобную регрессионному уравнению.

Необходимость в особом методе анализа, отличном от регрессионного, связана с тем, что последний не позволяет корректно учитывать информацию о случаях, для которых событие за время наблюдения не наступило. Такие случаи называют цензурированными справа. Кроме того, регрессионный анализ не допускает использования независимых переменных, изменяющихся во времени. Например, вероятность вступления в брак в разные периоды жизни может варьировать с изменениями дохода, жилищных условий и образования.

В зависимости от концептуализации событий различают три стратегии А.С.: для единичных неповторяющихся событий (вступление в первый брак); для множественных событий (вступление в одну из политических партий); для повторяющихся событий (переход с одной работы на другую).

А.С. является обобщением ряда методов, разработанных в частных науках для эксплораторного исследования единичных неповторяющихся событий: анализа отказов в технике, анализа дожития (выживания) в биологии, анализа таблиц дожития в демографии. Общая модель А.С. для непрерывного времени была разработана американскими социологами Н. Тума и М. Ханнаном в конце 1970-х и является одним из главных методологических достижений социологии за последние 25 лет. Ключевыми понятиями А.С. являются функция плотности вероятности $f(t)$ (безусловная вероятность того, что событие произойдет во время t), функция распределения накопленной вероятности

$F(t)$, функция дожития $S(t)$ (вероятность того, что событие не произойдет вплоть до момента времени t). Особое значение имеет интенсивность, буквально "показатель риска" (hazard rate), или вероятность наступления события в малом промежутке времени

между t_j и $t_j + \Delta t$, при условии, что событие не произошло до этого. Для непрерывно измеренного времени

.

По статистическим и содержательным соображениям социологи предпочитают исследование функциональной формы интенсивности исследованию других упомянутых функций. Интенсивность концептуально связана с другими понятиями А.С.: $h(t) = f(t)/S(t) = -d(\ln S(t))/dt$.

Как и другие статистические методы, А.С. может быть эксплораторным (описательным) либо конфирматорным (направленным на проверку гипотез). В первом случае строят и сравнивают эмпирические графики функций, например выживания, для различных значений независимых переменных, например для мужчин и женщин. Вторая стратегия состоит в моделировании функциональной формы зависимости интенсивности (риска) наступления события от истекшего времени и от значений предикторов. Цель в том, чтобы не только оценить эффекты независимых переменных, но и подобрать (на основании содержательной теории или эмпирически) наилучшую функциональную зависимость, т.е. модель, способную наилучшим образом объяснить гетерогенность совокупности.

И функция времени, и гетерогенность совокупности могут моделироваться с помощью двух подходов. Непараметрический подход связан с дискретизацией данных и оценкой показателя риска для каждого из значений независимых переменных. Предположений о форме функциональной зависимости не делается.

В параметрическом подходе функциональные формы зависимостей постулируются и оцениваются на предмет правдоподобия.

Чаще всего используют следующие функции зависимости от времени: экспоненциальную, Вейбулла, Гернеса, Мейкхама-Гомперца для монотонной связи интенсивности с временем, лог-логистическую, логарифмически-гауссову для немонотонной связи. Специальным случаем моделирования гетерогенности совокупности являются модель пропорционального риска, в которой интенсивность (показатель риска) есть функция базового показателя риска, $q(t)$, зависящего от времени и постоянного для всех единиц выборки (респондентов), и вектора предикторов, x , не зависящего от времени: $h(t, x) = q(t) \cdot e^{x\beta}$, где β - вектор регрессионных коэффициентов.

Для оценки этой модели Д. Кокс предложил метод частичного правдоподобия, который позволяет исследователям не делать предположений о $q(t)$, что удобно при отсутствии как содержательных теорий, так и специальных статистических знаний. С этим связана большая популярность регрессионных моделей Кокса.

А.С. требует сбора данных в специальном формате, с точным измерением времени наступления событий, что затрудняет применение метода на практике.

С некоторой потерей информации задачу можно переформулировать для дискретного времени, когда время предполагается измеренным с низкой точностью, и оценить обычную логистическую регрессию, в которой зависимой переменной является факт наступления или ненаступления события. Серьезные проблемы для анализа создает левое цензурирование, т.е. отсутствие информации о предыстории событий. Поскольку многие предикторы сами зависят от времени, для моделирования даже относительно простых событий (рождение ребенка) необходимо располагать полной историей жизни респондента, включая образовательную, трудовую и др. линии. Одной из актуальных задач статистического аппарата А.С. является включение в модели ошибок измерения, подобно тому, как это делается в структурных уравнениях.

А.С. продуктивно использовался для анализа брачно-репродуктивного поведения (Н. Тума и др.), влияния человеческого капитала на брачное поведение (Х.-П. Блоссфельд), образования и ликвидации фирм (М. Ханнан и Г. Кэррол), трудовых биографий (А. Соренсен), принятия законов разными штатами США (Э. Повалко), обретения независимости бывшими колониями (Д. Странг), этнических конфликтов (С. Ользак) и др. А.С. позволяет рассматривать процессы диффузии (распространения слухов или инноваций) на микроуровне, с учетом позиции индивидов в социальной структуре.

С.В. Сивуха