



ВЫБОРКА И ВЫБОРОЧНОЕ РАСПРЕДЕЛЕНИЕ

Зачем мы делаем выборку?

В статистике популяция представляет целую группу индивидуумов, которые нас интересуют. Вообще изучать целую популяцию довольно дорого и трудоемко, а в некоторых случаях просто невозможно, так как популяция может быть гипотетической (например, пациенты, которые будут проходить лечение в будущем), поэтому собирают данные по выборке индивидуумов, как предполагают, представителей этой популяции, и используют их для того, чтобы сделать выводы (т.е. делать заключения) относительно этой популяции.

Зачем мы делаем выборку?

Когда берут выборку из популяции, имеют ввиду, что информация в выборке не может полностью отражать то, что истинно в этой популяции. Возможна ошибка, обусловленная выборкой, так как изучалась только лишь часть популяции





Получение репрезентативной (представительной выборки)

В идеале мы стремимся к *случайной (рандомизированной) выборке*. Составляют список всех индивидуумов популяции (структура выборки) и из этого списка случайно отбирают индивидуумов, т.е. каждая возможная выборка данного объема имеет равную вероятность быть выбранной из популяции.

Иногда возникают трудности при составлении этого списка или материальные ограничения, и в этом случае берут приемлемую выборку.

ТОЧЕЧНЫЕ ОЦЕНКИ

Мы часто заинтересованы в оценке параметра в популяции, среднего или пропорции (доли). Параметры обычно обозначают символами греческого алфавита. Например: обычно мы обозначаем среднее популяции как μ , а стандартное отклонение популяции как σ .

Мы оцениваем значение параметра, используя данные, собранные в выборке. Эта оценка – точечная оценка генерального параметра (т.е. она принимает только одно значение) в отличие от интервальной оценки, которая имеет интервал значений. Точечную оценку описывает *выборочная статистика*.

Выборочная дисперсия

Если повторить извлечение выборок того же самого объема из популяции, маловероятно, что оценки параметра популяции будут точно такими же в каждой выборке. Однако все оценки должны быть близки к истинному значению параметра (генеральному параметру) в популяции и подобны друг другу.

Определяя величину вариабельности этих оценок, мы поймем, насколько они точны, и таким образом сможем оценить ошибку, обусловленную выборкой. Обычно берут только одну выборку из популяции.

Выборочное распределение среднего

Предположим, что мы заинтересованы в оценке среднего популяции; можно брать много повторных выборок объема n из популяции и оценить среднее в каждой выборке. Гистограмма оценок этих средних показала бы их распределение; это – распределение выборочных средних. Мы можем увидеть следующее:

- ❖ если объем выборки разумно большой, оценки среднего имеют нормальное распределение при любом распределении исходных данных в популяции (это следует из теоремы, известной как центральная предельная теорема);

Выборочное распределение среднего

- ❖ если объем выборки небольшой, оценки среднего отвечают нормальному распределению при условии, что данные в популяции также отвечают нормальному распределению;
- ❖ среднее этих оценок – несмещенная оценка истинного среднего в популяции (генерального среднего), т.е. среднее этих оценок эквивалентно истинному среднему в популяции;
- ❖ Вариабельно распределения выражается стандартным отклонением оценок, известным как стандартная ошибка среднего (часто обозначают как Standard Error Means, SEM). Если бы мы знали стандартное отклонение популяции σ , тогда стандартная ошибка среднего описывалась бы так...

Выборочное распределение среднего

- ❖ В случае если есть, как обычно, только одна выборка, нашей лучшей оценкой среднего популяции будет выборочное среднее, а так как редко бывает известно стандартное отклонение в популяции (генеральный стандарт), то стандартную ошибку среднего оценивают следующим образом...
Стандартная ошибка среднего отражает точность нашей оценки.

Интерпретация стандартной ошибки

- ❖ Большая стандартная ошибка указывает, что оценка неточна.
- ❖ Небольшая стандартная ошибка указывает, что оценка точна.
- ❖ Стандартная ошибка уменьшится, т.е. мы получим более точную оценку, если:
 - ✓ объем выборки увеличится;
 - ✓ данные имеют небольшое рассеяние.

Стандартное отклонение (SD) или стандартная ошибка среднего (SEM)

Хотя на первый взгляд кажется, что эти два параметра схожи, их используют в разных целях. Стандартное отклонение отражает вариативность в значениях данных и должно быть указано, если нужно пояснить изменчивость в наборе данных. Наоборот, стандартная ошибка отображает точность выборочного среднего и должна быть указана, если интересует среднее значение набора данных.

ПРОВЕРКА ГИПОТЕЗ

Часто собирают выборочный набор данных для того, чтобы оценить, сколько существует аргументов против определенной гипотезы, сформулированной относительно популяции (генеральной совокупности). Используют процесс, известный как проверка гипотез (или проверка значимости), чтобы определить количественную меру уверенности против определенной гипотезы.

Проверка гипотез. Общий обзор.

- I. Определить нулевую и альтернативную гипотезу при исследовании.
- II. Отобрать необходимые данные из выборки пациентов.
- III. Вычислить значение статистики критерия, отвечающий H_0 .
- IV. Сравнить значение статистики критерия со значениями из известного распределения вероятности (для данной статистики).
- V. Интерпретировать достигнутый уровень значимости P -значение.

Определение нулевой и альтернативной гипотезы

Мы всегда проверяем нулевую гипотезу (H_0), которая отвергает эффект (к примеру, разница в средних равна 0) в популяции.

Затем мы определяем альтернативную гипотезу (H_1), которая принимается, если нулевая гипотеза не верна.

Альтернативная гипотеза в большей степени относится к той теории, которую мы хотели бы исследовать.

ПОЛУЧЕНИЕ СТАТИСТИКИ КРИТЕРИЯ

Когда данные собраны, подставляем значения из выборки в формулу, используемую для критерия, который применяют, чтобы определить значение статистики критерия. Это величина количественно отражает аргументы в наборе данных против нулевой гипотезы — обычно чем больше значение статистики критерия, игнорируя его знак, тем сильнее эти аргументы.

ПОЛУЧЕНИЕ P-ЗНАЧЕНИЯ

Все статистики критерия подчиняются известным теоретическим распределениям вероятности. Мы связываем значение статистики критерия, полученное из выборки с уже известным распределением, которому она подчиняется, чтобы получить P-значения, площадь обоих (или иногда одного) «хвостов» распределения вероятности. Большинство компьютерных пакетов обеспечивают автоматическое вычисление двустороннего P-значения.

ПОЛУЧЕНИЕ P-ЗНАЧЕНИЯ

P-значение есть вероятность получения нашего вычисленного значения критерия или еще большего значения, если нулевая гипотеза истинна. Нулевая гипотеза всегда относится к популяции, представляющей для нас интерес больший, нежели выборка. Следовательно, нулевая гипотеза может быть либо истинной, либо ложной, и мы не можем интерпретировать P-значение как вероятность того, что нулевая гипотеза истинна.

Применение Р-значения

Мы должны принять решение относительно того, сколько аргументов потребуется для того, чтобы была возможность решить, отвергнуть ли нулевую гипотезу в пользу альтернативной.

Чем меньше Р-значение, тем сильнее аргументы против нулевой гипотезы.

Применение Р-значения

Традиционно мы полагаем, что если Р-значение меньше 0,05, то имеется достаточно аргументов, для того, чтобы отвергнуть нулевую гипотезу, хотя есть небольшой шанс появления результатов, если нулевая гипотеза истинна.

Тогда мы отвергнем нулевую гипотезу и скажем, что результаты значимы на 5% уровне.

Применение Р-значения

Если $P > 0,05$, то обычно мы приходим к выводу, что аргументов недостаточно для того, чтобы отвергнуть нулевую гипотезу. Мы не отвергаем нулевую гипотезу и можем заявить, что результаты не значимы на 5% уровне.

Это совсем не означает, что нулевая гипотеза истинна; просто у нас нет достаточных аргументов, чтобы отвергнуть ее.

Применение Р-значения

Выбор 5% произволен. Если брать за основу 5%, то мы поступим неправильно, отвергнув нулевую гипотезу, когда она истинна. Неправильное отклонение нулевой гипотезы может привести к серьезным последствиям, в этом случае необходимо потребовать более веских аргументов, прежде чем отвергнуть нулевую гипотезу, например, мы можем выбрать Р-значение 0,01 (или 0,001).

Выбранную границу отсечки (например, 0,05 или 0,01) называют критическим уровнем значимости критерия.

Применение P-значения

Определение результата только как значимого на определенном уровне граничного значения (например, $P < 0,05$) может ввести в заблуждение.

Допустим, если $P = 0,04$, то мы отвергаем нулевую гипотезу; однако если $P = 0,06$, то мы не отвергнем ее.

Действительно ли они различны?

Мы рекомендуем всегда указывать точное P-значение, обычно получаемое из результата компьютерного анализа.