

Методические указания для выполнения индивидуальных заданий.

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

- *Статистической гипотезой* называется предположение о виде неизвестного распределения случайной величины или о параметрах известного распределения. Наряду с проверяемой гипотезой (*нулевой, или основной*) H_0 формулируется и противоречащая ей гипотеза (*конкурирующая, или альтернативная*) H_1 , которая принимается, если отвергнута нулевая гипотеза.

Гипотезы разделяются

- Гипотезы разделяются на *простые* (содержащие только одно предположение) и *сложные* (содержащие более одного предположения).
- При проверке гипотезы могут быть допущены ошибки двух видов: *ошибка первого рода*, если отклонена верная нулевая гипотеза, и *ошибка второго рода*, если принята неверная нулевая гипотеза.

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

- Для проверки статистической гипотезы используется специально подобранная случайная величина K с известным законом распределения, называемая *статистическим критерием*. Множество ее возможных значений разбивается на два непересекающихся подмножества: одно из них (*критическая область*) содержит значения критерия, при которых нулевая гипотеза отклоняется, второе (*область принятия гипотезы*) – значения K , при которых она принимается. Значения K , отделяющие критическую область от области принятия гипотезы, называются *критическими точками* k_p .

- Например, при сравнении показателей курения у мужчин и женщин в популяции нулевая гипотеза означала бы, что показатели курения одинаковые у женщин и мужчин в популяции.
- Затем определяют альтернативную гипотезу (H_1), которая принимается, если нулевая гипотеза неверна. Альтернативная гипотеза больше относится к той теории, которую собираются исследовать. Итак, на этом примере альтернативная гипотеза заключается в утверждении, что показатели курения различны у женщин и мужчин в популяции.
- Разницу в показателях курения не уточнили, т.е. не установили, имеют ли в популяции мужчины более высокие или более низкие показатели, чем женщины. Такой подход известен как *двусторонний критерий*, потому что учитывают любую возможность, он рекомендуется постольку, поскольку редко есть уверенность заранее в направлении какого-либо различия, если таковое существует.

Двусторонняя критическая область

- В некоторых случаях можно использовать односторонний критерий для гипотезы , в котором направление эффекта задано. Его можно применить, например, если рассматривать заболевание, от которого умерли все пациенты, не получившие лечения; новый препарат не мог бы ухудшить положение дел.

Левосторонняя критическая область

Правосторонняя критическая область

- Критическая область может быть *правосторонней* (если она задается неравенством

$$K > k_{кр}$$

- *левосторонней*

$$K < k_{кр}$$

- *Двусторонней*

$$K < (k_{кр})_1, \quad K > (k_{кр})_2$$

-

Если критическая область симметрична относительно 0, то двусторонняя критическая область определяется:

Нахождение правосторонней критической области

- Для ее нахождения нужно задать вероятность ошибки первого рода α , называемую *уровнем значимости*; тогда, например, правосторонняя критическая область задается условием

$$p(K > k_{кр}) = \alpha$$

Со сложными понятиями легче разобратся на примере.

- Во время производства некоего лекарства от учёных требуется чрезвычайная осторожность, так как превышение дозы одного из компонентов провоцирует высокий уровень токсичности готового препарата, от которого пациенты, принимающие его, могут умереть. Однако на химическом уровне выявить передозировку невозможно. Из-за этого перед тем как выпустить лекарство в продажу, небольшую его дозу проверяют на крысах или кроликах, вводя им препарат. Если большая часть испытуемых умирает, то лекарство в продажу не допускается, если подопытные живы, то лекарство разрешают продавать в аптеках.

- Первый случай: на самом деле лекарство было не токсично, но во время эксперимента была допущена оплошность и препарат классифицировали как токсичный и не допустили в продажу. $A=1$.
- Второй случай: в ходе другого эксперимента при проверке другой партии лекарства решено, что препарат не токсичен, и в продажу его допустили, хотя на самом деле препарат был ядовит. $A=2$.
- Первый вариант повлечёт за собой крупные финансовые затраты поставщика-предпринимателя, так как придётся уничтожить всю партию лекарства и начинать с нуля.
- Вторая ситуация спровоцирует смерть пациентов, купивших и употреблявших это лекарство.

Порядок проверки статистической гипотезы таков:

- задается уровень значимости α , выбирается статистический критерий K и вычисляется (обычно по таблицам для закона распределения K) значение $k_{кр}$; определяется вид критической области;
- по выборке вычисляется наблюдаемое значение критерия $K_{набл}$;
- если $K_{набл}$ попадает в критическую область, нулевая гипотеза отвергается; при попадании $K_{набл}$ в область принятия гипотезы нулевая гипотеза принимается.

Проверка гипотезы о значимости выборочного коэффициента корреляции

- По выборке объема $n = 150$, извлеченной из нормально распределенной двумерной генеральной совокупности, вычислен выборочный коэффициент корреляции
- $(r_B) = - 0,37$. Проверим при уровне значимости
- $\alpha = 0,01$ нулевую гипотезу H_0 :
- $P(r_{r'}) = 0$ о равенстве нулю генерального
- коэффициента корреляции при конкурирующей гипотезе H_1 :
 $r_{r'} \neq 0$.

Решение.

Сравнение двух дисперсий нормальных генеральных совокупностей

- Пусть имеются две выборки объемов n_1 и n_2 , извлеченные из нормально распределенных генеральных совокупностей X и Y . Требуется по исправленным выборочным дисперсиям и проверить нулевую гипотезу о равенстве генеральных дисперсий рассматриваемых генеральных совокупностей

- Величина r является показателем тесноты линейной связи и называется выборочным коэффициентом корреляции (или просто коэффициентом корреляции).