

## Проверка статистических гипотез.

выборка  $\{X_1, X_2, \dots, X_n\}$  из некоторой генеральной совокупности

**Определение 2.** *Критерий* — это решающее правило, по которому принимают или отвергают гипотезы, обеспечивающее надежное поведение, то есть принятие истинной и отклонение ложной гипотезы с высокой вероятностью.

**Определение 3.**

$P(H_1 | H_0) = \alpha$  - ошибка первого рода (уровень значимости) - вероятность отвергнуть гипотезу  $H_0$ , если она верна (то есть принять гипотезу  $H_1$ ).

$P(H_0 | H_0) = 1 - \alpha = \gamma$  - уровень доверия (надежности) - вероятность принять верную гипотезу.

$P(H_0 | H_1) = \beta$  - ошибка второго рода - вероятность принять основную гипотезу если она неверна.

$P(H_1 | H_1) = 1 - \beta$  - мощность критерия - вероятность принять гипотезу  $H_1$ , если она верна.

## Проверка статистических гипотез.

**Пример.** Строительная компания хочет построить дом, она обращается в комиссию, где должны принять решение о том, можно строить дом или нельзя.

Основная гипотеза  $H_0$  - решение о том, что нельзя строить дом на определенном участке,

альтернативная гипотеза  $H_1$  - решение о том, что дом строить можно.

Ошибка первого рода  $P(H_1 | H_0) = \alpha$  – принятие решения о том, что дом строить можно, хотя на самом деле этого делать нельзя.

Ошибка второго рода  $P(H_0 | H_1) = \beta$  – принятие решения о том, что дом строить нельзя, хотя на самом деле это делать можно.

Хотелось бы уменьшить обе ошибки, но это, к сожалению, невозможно.

## Проверка статистических гипотез.

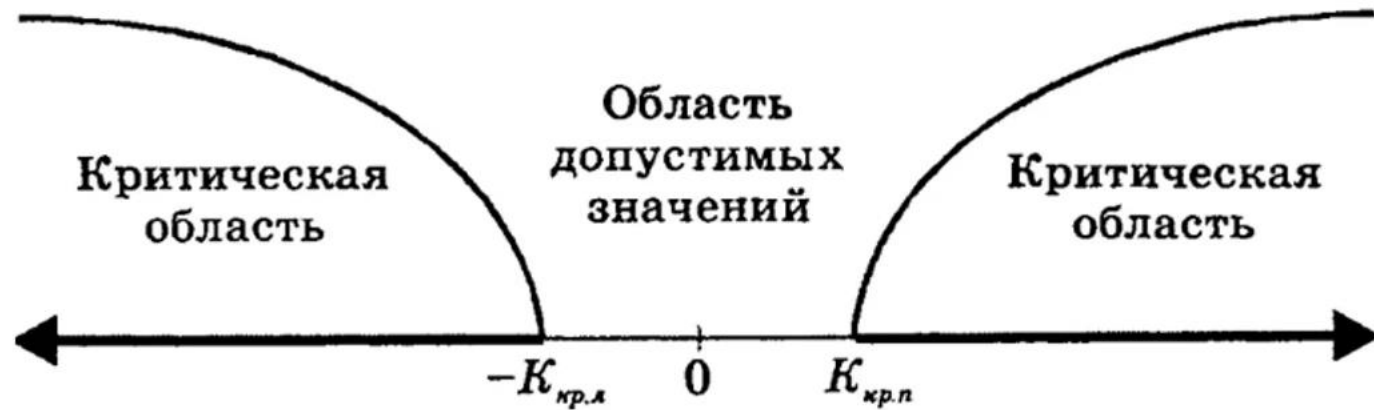
**Определение 4.** *Статистический критерий* – случайная величина, имеющая определенное распределение в случае, если верна гипотеза  $H_0$ .

1. Выбирается критерий  $K$  так, чтобы в случае, если гипотеза  $H_0$  верна, распределение случайной величины  $K = K(X_1, \dots, X_n)$  было бы известным.
2. Вычисление эмпирического значения критерия (подстановка данных в формулу)
3. Выбор критических значений  $K_1$  и  $K_2$  (по таблице)- зависят от ошибки первого рода
4. Сравнение критических и эмпирического значений – вывод.

## Проверка статистических гипотез.

**Определение 5.** *Критическая область* – если в эту область попало эмпирическое значение, то мы отвергаем  $H_0$ , а значит, принимаем  $H_1$ .

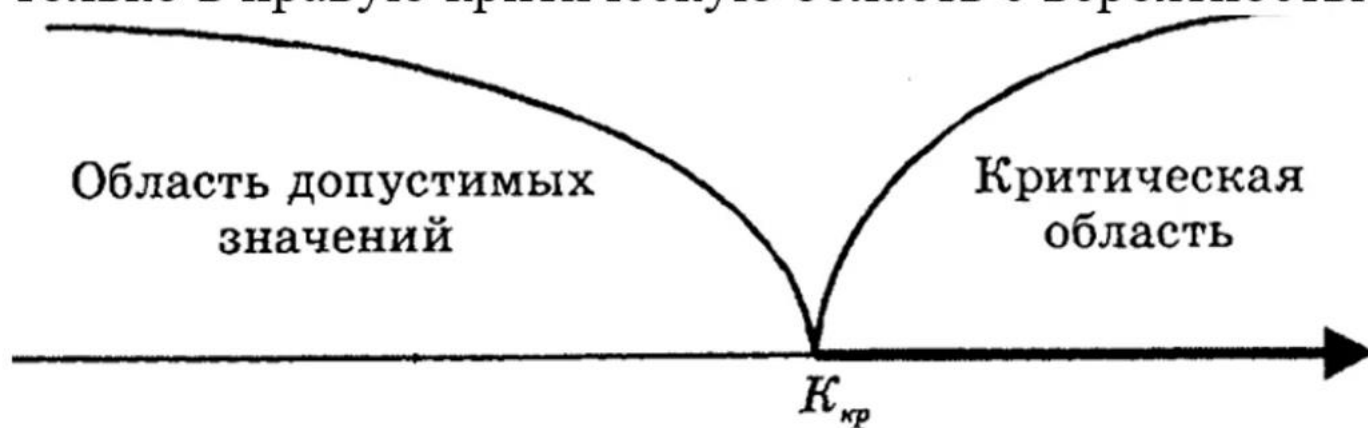
а) Двусторонняя критическая область. Существует вероятность ошибиться – попасть и в левую, и в правую критические области, в каждую с вероятностью  $\alpha/2$ .



## Проверка статистических гипотез.

**Определение 5.** *Критическая область* – если в эту область попало эмпирическое значение, то мы отвергаем  $H_0$ , а значит, принимаем  $H_1$ .

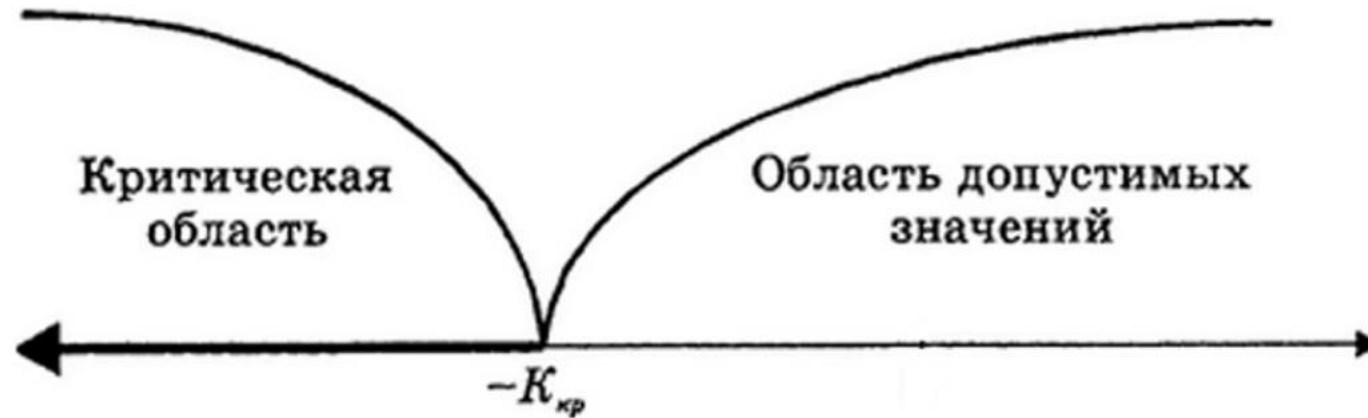
б) Правосторонняя критическая область. Существует вероятность ошибиться – попасть только в правую критическую область с вероятностью  $\alpha$ .



## Проверка статистических гипотез.

**Определение 5.** *Критическая область* – если в эту область попало эмпирическое значение, то мы отвергаем  $H_0$ , а значит, принимаем  $H_1$ .

в) Левосторонняя критическая область. Существует вероятность ошибиться – попасть только в левую критическую область с вероятностью  $\alpha$ .



## Проверка статистических гипотез.

**Определение 6.** *Параметрический* критерий – критерий, в формуле которого присутствуют оценки параметров распределения – выборочное среднее и/или выборочная дисперсия. Применяются только для количественных признаков.

**Определение 7.** *Непараметрический* критерий – критерий, используемый в тех случаях, когда исследователь ничего не знает о параметрах исследуемого признака. Применяются для признаков любого типа.