

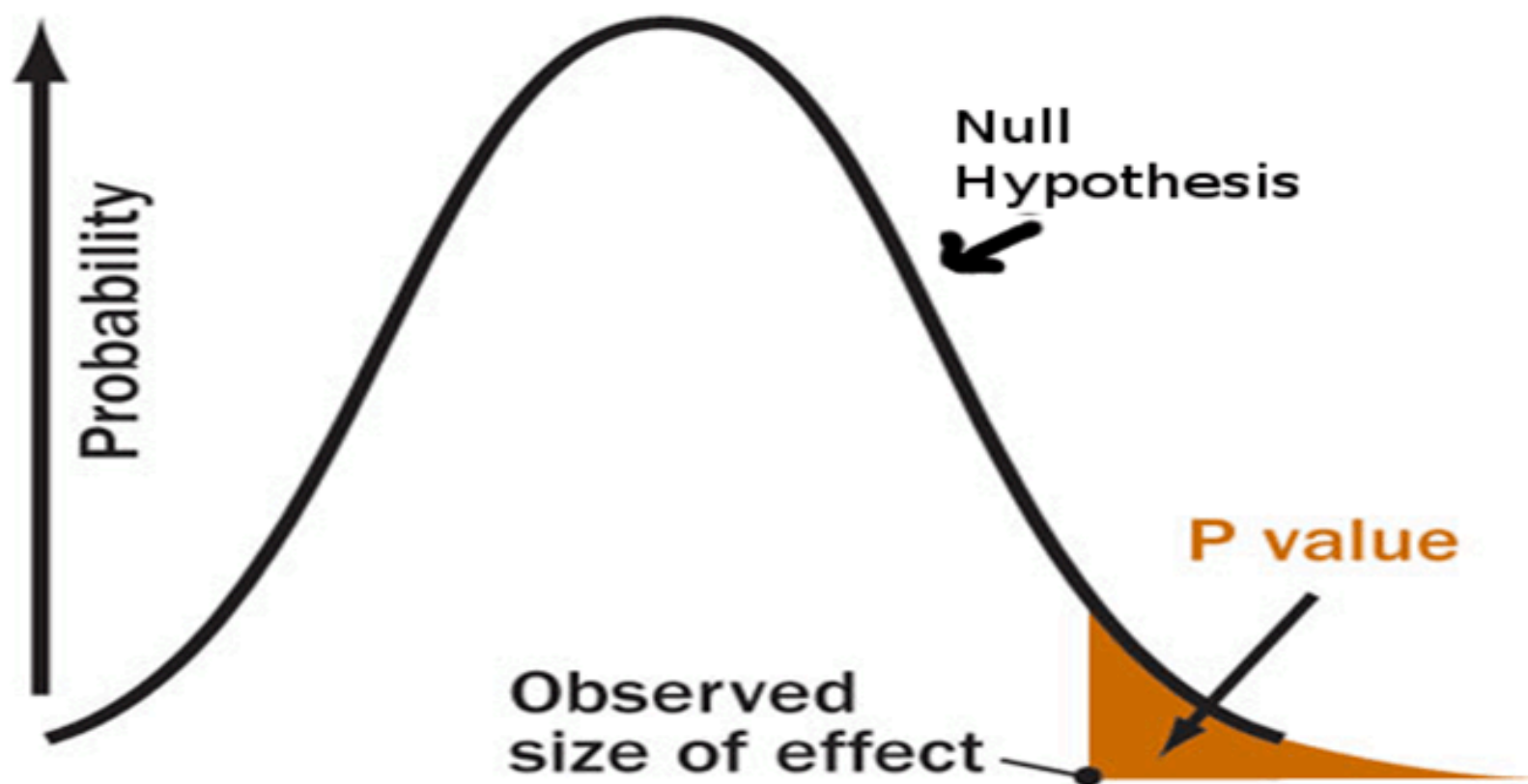
Статистическая значимость. Методы проверки гипотез

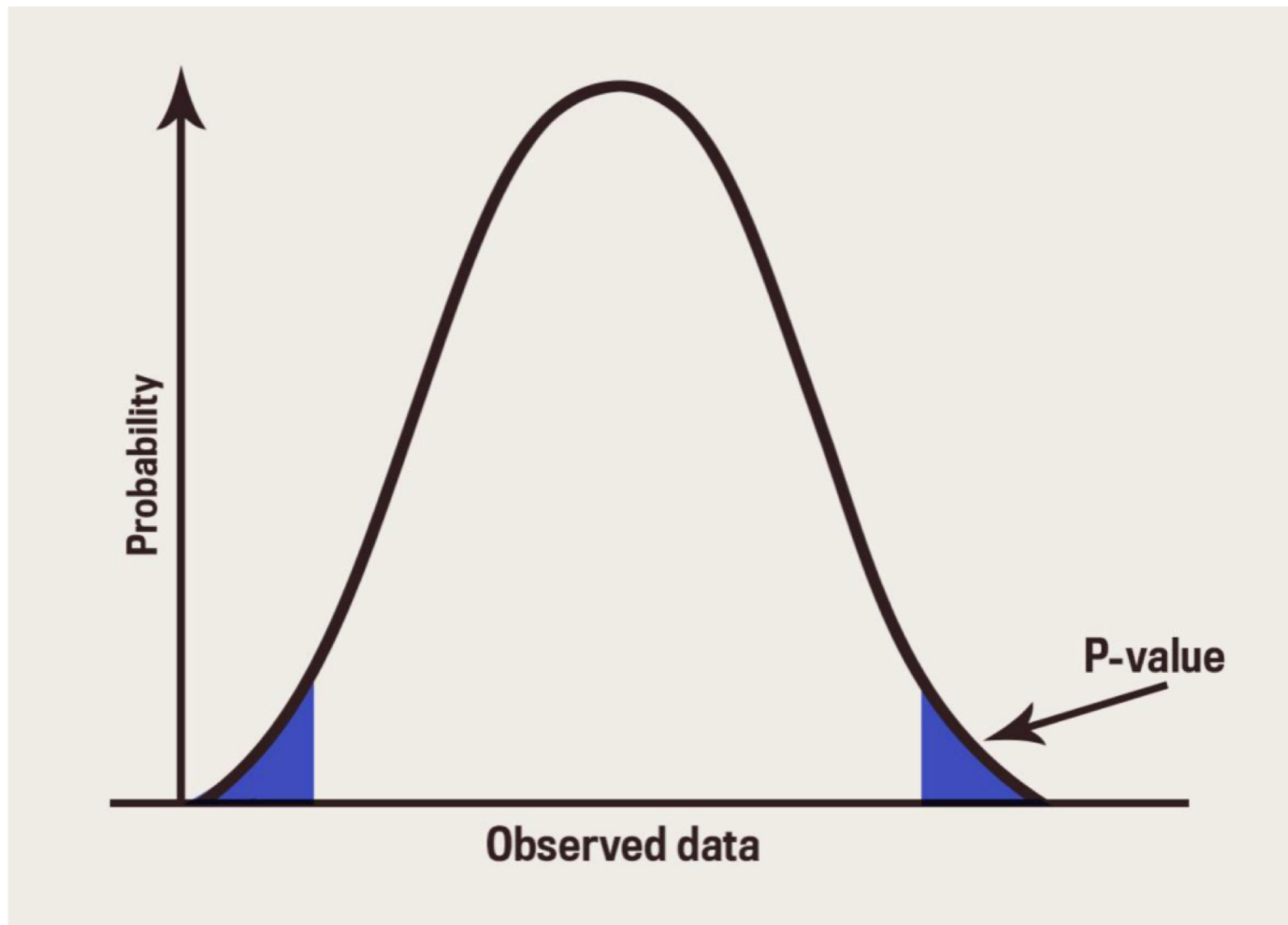
Величина называется **статистически значимой**, если её (или более крайних величин) случайное появление маловероятно.

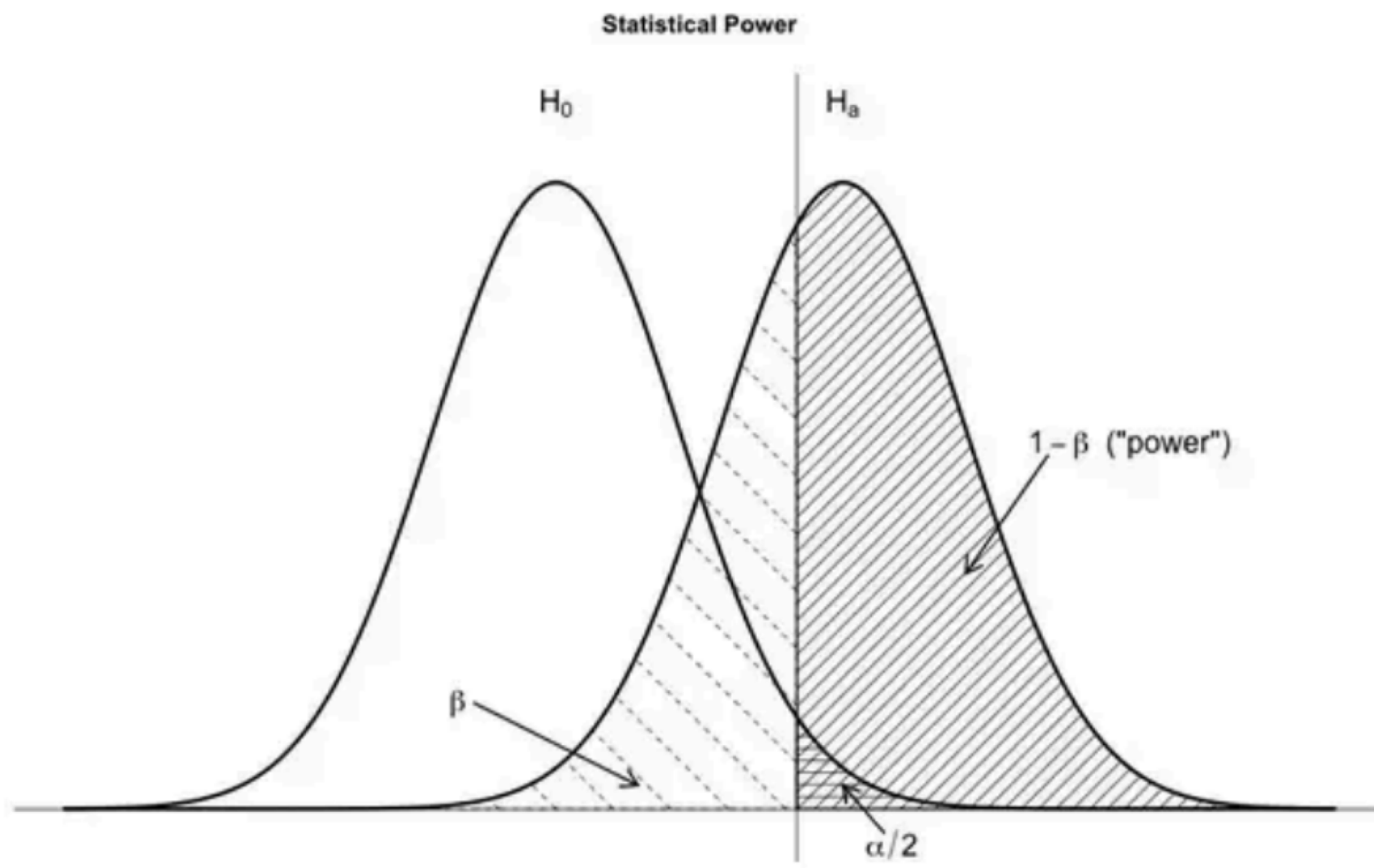
Нулевая гипотеза - принимаемое по умолчанию предположение о том, что не существует связи между двумя наблюдаемыми событиями.

Альтернативная гипотеза утверждает наоборот, что связь существует. Бывает **односторонней** и **двусторонней**.

P-value - вероятность получить для данной вероятностной модели распределения значений случайной величины такое же или более экстремальное значение статистики (среднего арифметического, медианы и др.), по сравнению с ранее наблюдаемым, при условии, что **нулевая гипотеза** верна.







Статистический критерий – правило, в соответствии с которым принимается или отвергается статистическая гипотеза.

Бывают **параметрические** и **непараметрические**.

t-критерий Стьюдента – применяется на проверку равенства средних двух выборок или одной с мат.ожиданием. *Является параметрическим критерием.*

$$t = \frac{\overline{X} - m}{s_X / \sqrt{n}}$$

Одновыборочный критерий

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Двухвыборочный критерий

U-критерий Манна — Уитни – критерий, используемый для оценки различий между двумя независимыми выборками по уровню какого-либо признака, измеренного количественно.

Позволяет выявлять различия в значении параметра между малыми выборками.

Является непараметрическим критерием.

$$U_1 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - R_2$$

$$U_1 + U_2 = n_1 \cdot n_2$$