11. О проверке статистических гипотез

Гипотеза **(H) – предположение о свойствах совокупности, проверяемое по выборочным данным**

(о распределении случайной величины, об его виде или параметрах)

При проверке по выборке $(x_1, x_2, \ldots, x_n)$ «выясняется», является ли ***отклонение*** от гипотезы ***случайным*** – тогда она считается ***верной*** (***принимается***),

или отклонение нельзя считать случайным, оно ***значимо*** – тогда гипотеза ***отвергается***, считается ***неверной***

Н у л е в а я г и п о т е з а **H_0 – отклонение от которой приписывается случаю («*нуль-гипотеза*»)**

Альтернативная гипотеза **– H_1**

<u>Например:</u>

H_0: $\sigma_1^2 = \sigma_2^2$ (о равенстве дисперсий в 2-х ГС, об их однородности)

против H_1: $\sigma_1^2 \neq \sigma_2^2$

(дисперсии разные, совокупности не однородны)

Гипотезы проверяются:

♣ в предположении правильности H_0

♥ по некоторому *критерию* $Cr\,(x_1, x_2, \dots, x_n / H_0)$ – распределению подходящей статистики – функции выборочных значений

♠ при заданной вероятности ошибки – *риске* α

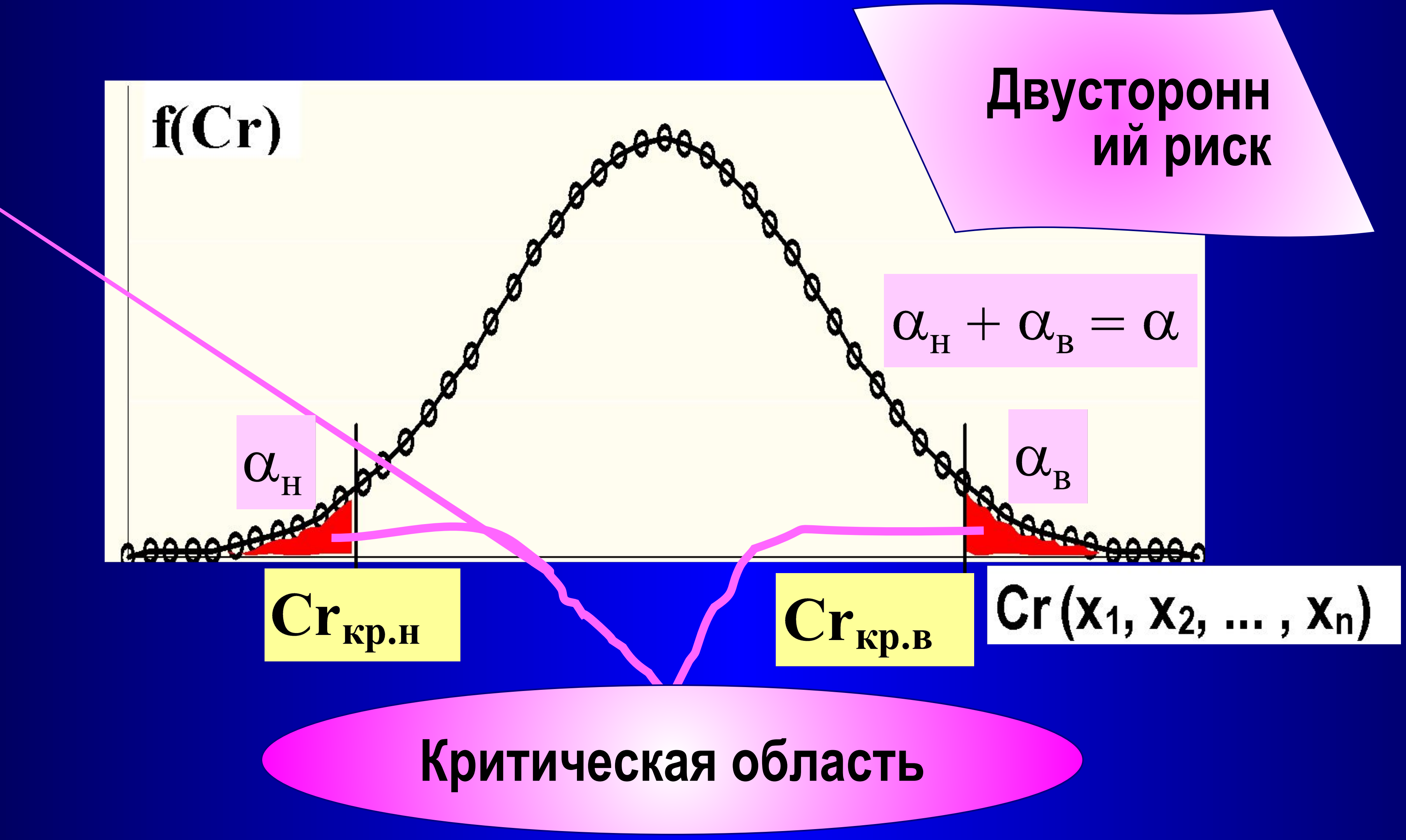
В примере про H_0: $\sigma_1^2 = \sigma_2^2$

используют «F-распределение» – ему подчиняется статистика $F = s_1^2 / s_2^2$ $(s_1^2 > s_2^2 \rightarrow F > 1)$ при числах степеней свободы $f_1 = n_1 - 1$ и $f_2 = n_2 - 1$, если справедлива H_0

Решающее правило

Если рассчитанное по $x_1, x_2, \ldots, x_n$ фактическое значение критерия ($Cr_ф$) попадает в область правдоподобных значений, то гипотеза допускается

если в *критическую область* неправдоподобно малых значений (ограниченных сверху критическим значением $Cr_{кр.н}$) или неправдоподобно больших (ограниченных снизу $Cr_{кр.в}$) – гипотеза отклоняется


Двусторонний риск
f(Cr)
$\alpha_н + \alpha_в = \alpha$
$\alpha_н$
$\alpha_в$
$Cr_{кр.н}$
$Cr_{кр.в}$
$Cr(x_1, x_2, \ldots, x_n)$
Критическая область

Односторонний риск

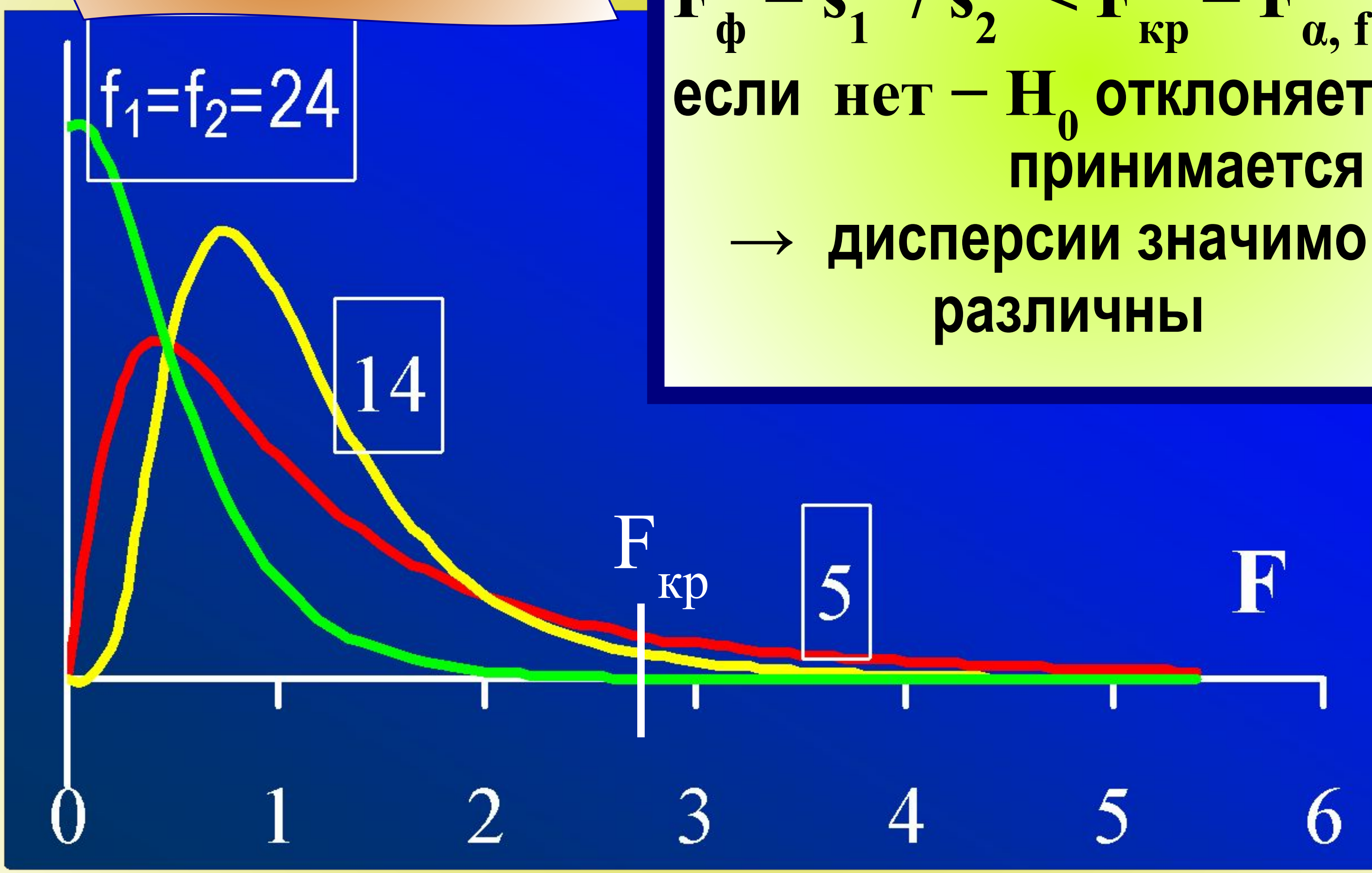
В случае с $H_0: \sigma_1^2 = \sigma_2^2$

она принимается, если

$F_ф = s_1^2 / s_2^2 < F_{кр} = F_{\alpha, f1, f2}$;

если нет – H_0 отклоняется, принимается H_1

→ дисперсии значимо различны



Статистический смысл решающего правила при проверке гипотез

При допущении истинности гипотезы, вероятность того, что значения критерия окажутся в критической области, мала (равна уровню значимости α) – это практически невозможное событие

Поэтому, если все же это оказывается так, логичен вывод, что допущение неверно, и гипотеза отклоняется.
Очевидно, что в α% случаев возможна ошибка – отклонение верной гипотезы

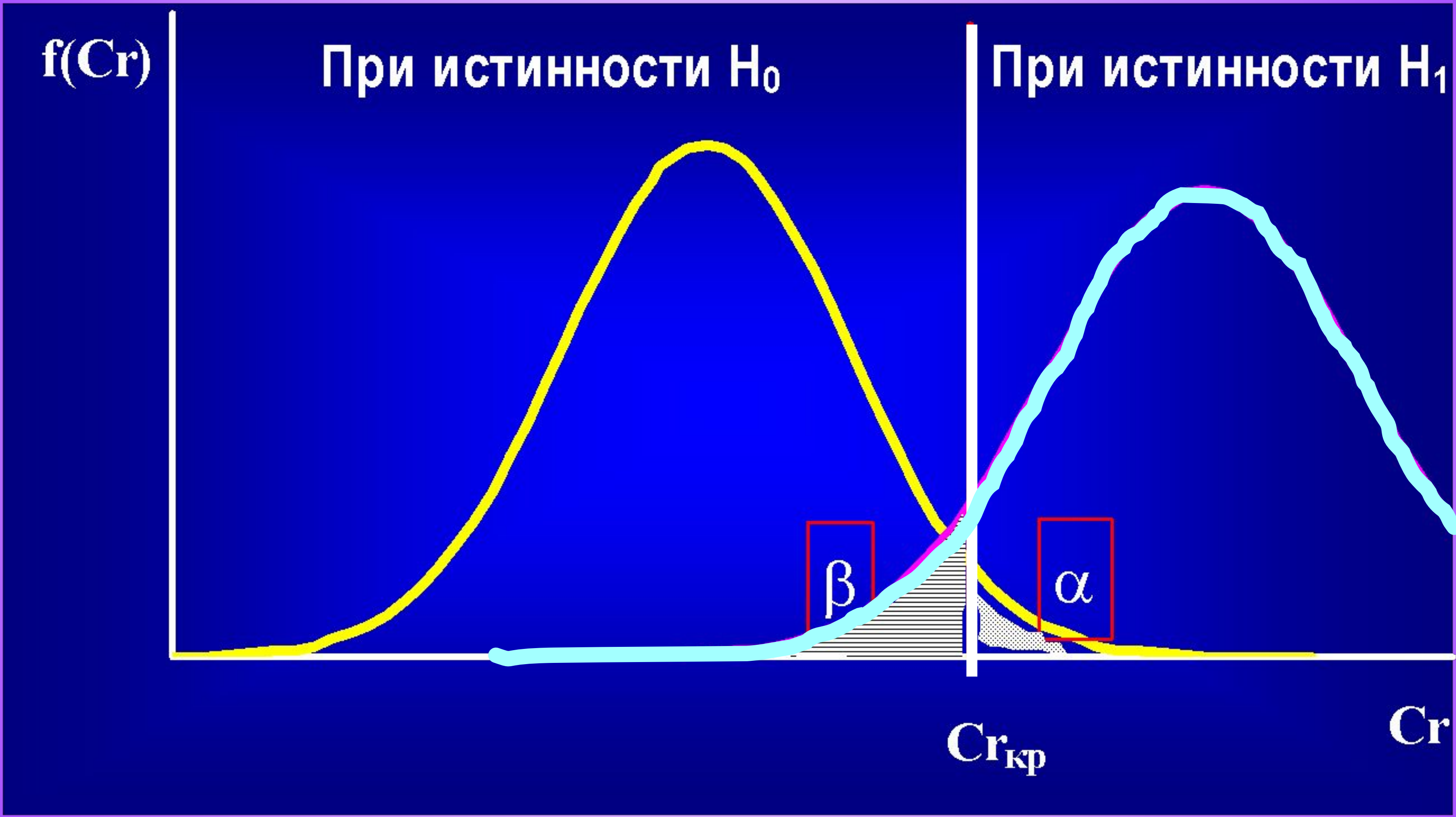
Возможны 2 рода ошибок в выводах

Ошибка 1-го

Ее вероятность α – «риск производителя»
отвергнуть партию годной продукции:
H_0 отвергается,
хотя является верной

Ошибка 2-го рода

Ее вероятность β – «риск потребителя»
принять негодную партию товара:
H_0 принимается,
хотя и неверна, а верна H_1



f(Cr)
При истинности H0
При истинности H1
β
α
Cr
Crкр

Таблица ситуаций при проверке гипотез

Фактически	Критерий рекомендует: Допустить H_0	Критерий рекомендует: Отклонить H_0, допустить H_1
Истинна H_0	Решение правдоподобно	Решение ложно – **ошибка 1-го рода**
Истинна H_1	Решение ложно – **ошибка 2-го рода**	Решение правдоподобно

Основные шаги при проверке гипотез

0. Формулировка H_0 и H_1

1. Подбор соответствующей статистики и ее распределения – статистического критерия Cr (теста)

2. Выбор риска α

3. Определение критической области ($Cr_{кр} = Cr_{\alpha}$)

4. Расчет по выборочным данным фактического значения критерия ($Cr_{ф}$)

5. Если оно попадает в критическую область, H_0 отклоняется, принимается H_1; если в область допустимых значений – H_0 не отклоняется

Пример

Задача на «Сравнение двух средних»

Деловая постановка задачи

Необходимо принять решение − какую из 2-х рекламируемых добавок следует закупать для производства фирменного материала: 2-ая дороже, но дает (судя по рекламе) **материал лучшего качества, что окупило бы дополнительные вложения.**

Для принятия решения требуется проверить (по результатам испытаний материала с каждой из добавок), **что 2-ая добавка действительно обеспечивает более высокий уровень качества (R)**

По результатам испытаний 6 образцов

с добавкой 1

$$\bar{R}_1 = 20.88$$

$$\mathbf{s_1 = 2.02}$$

$$\mathbf{n = 6 \quad f_1 = 5}$$

с добавкой 2

$$\bar{R}_2 = 23.38$$

$$\mathbf{s_2 = 1.67}$$

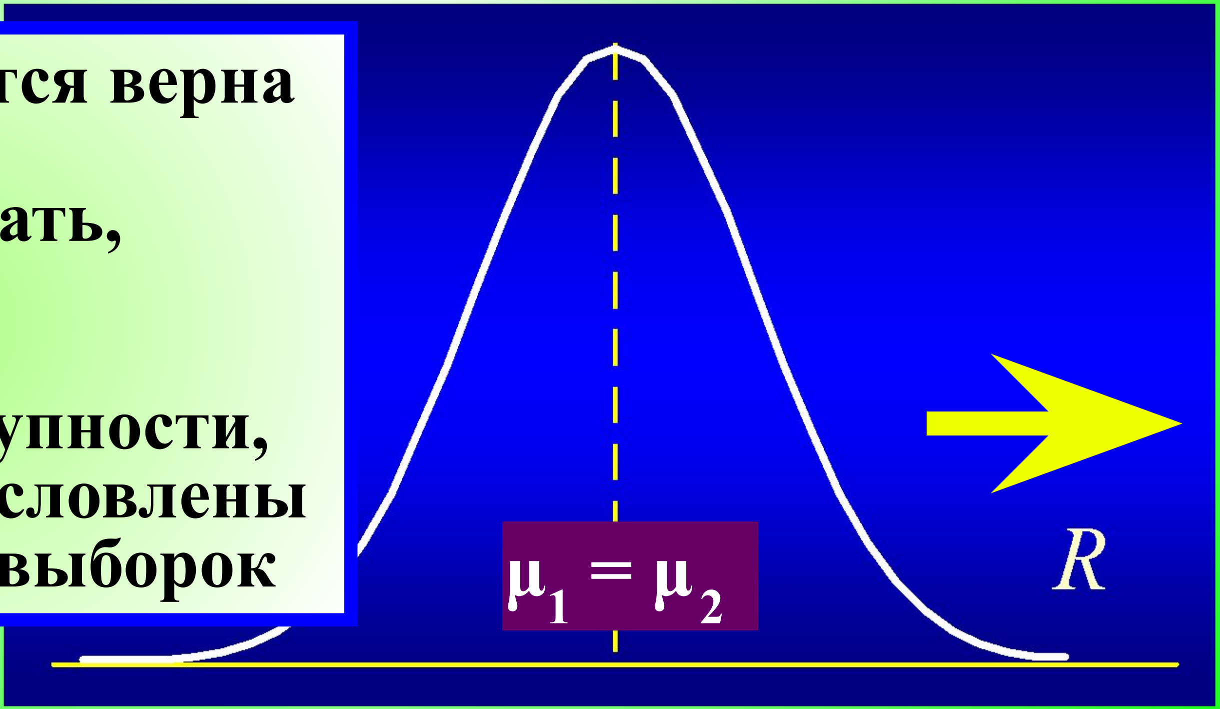
$$\mathbf{n = 6 \quad f_2 = 5}$$

Это не одно и то же ?

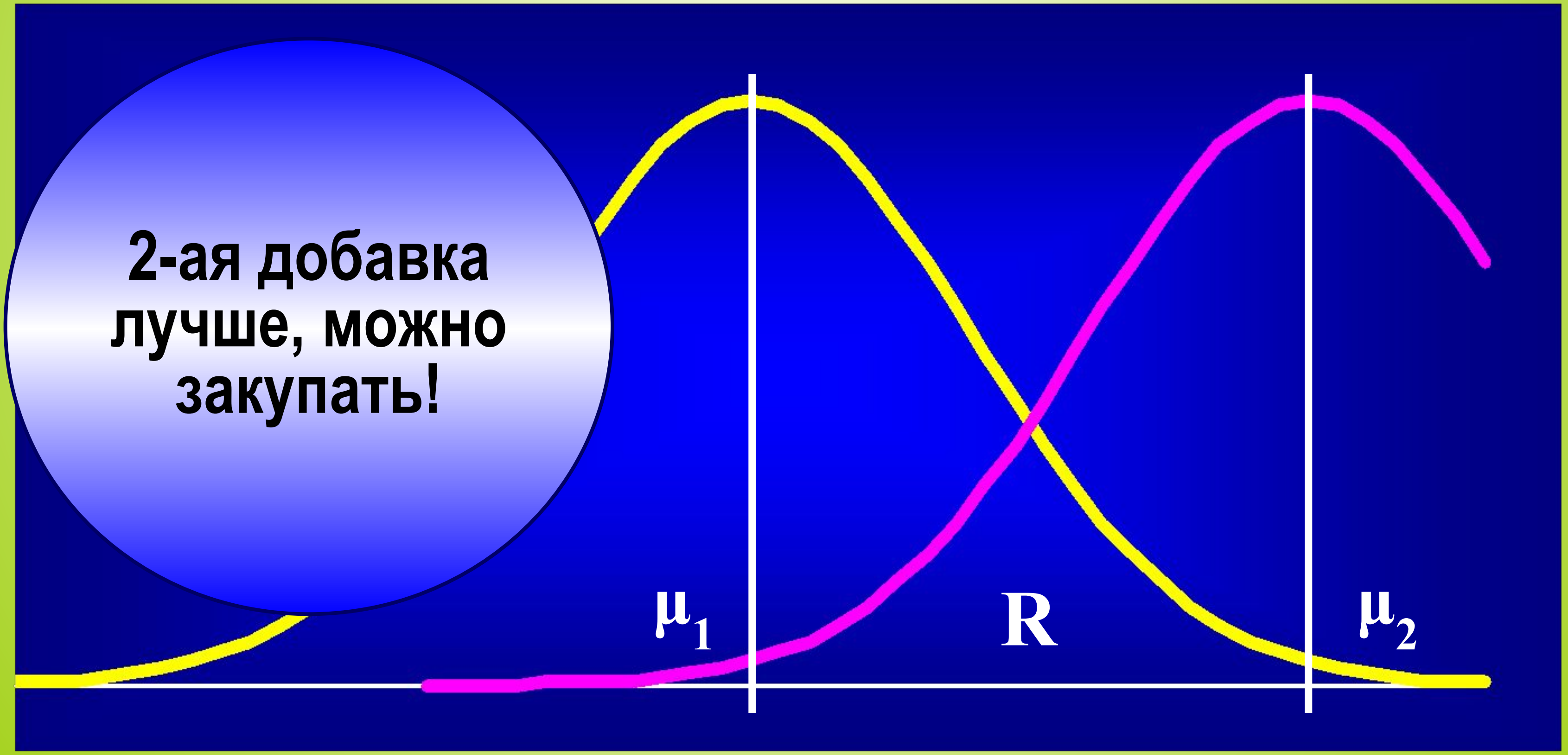
Статистическая постановка задачи

Деловая задача сводится к задаче о проверке гипотезы
$H_0: \mu_1 = \mu_2$, где μ – генеральное (истинное) среднее значений R
против альтернативной
$H_1: \mu_1 \neq \mu_2$ (или $H_1: \mu_2 > \mu_1$)

**Если H_0 окажется верна (принята),
это будет означать,
что R_1 и R_2 ,
на самом деле,
из одной совокупности,
а различия обусловлены
случайностью выборок**



Если H_0 будет отвергнута, то считаем, что верна альтернативная гипотеза H_1, и различия не случайны, значимы, выборки R_1 и R_2 из разных совокупностей



Проверка H_0 осуществляется по t-критерию, по которому распределена статистика

$$\tilde{t} = \frac{\left|\bar{R}_1 - \bar{R}_2\right|}{\sqrt{2\tilde{s}(\bar{R})}}$$

здесь $\tilde{s}^2(\bar{R})$ – обобщенная дисперсия средних при соответствующем числе степеней свободы; определяется после проверки гипотезы об однородности σ_1 и σ_2 ;

таким образом, в знаменателе оценка σ разности →

«смысл» t – во сколько раз различие в средних больше меры их случайного рассеяния

В нашей задаче – по результатам испытаний 6 образцов

$$t = \frac{|20.88 - 23.38|}{\sqrt{2 \cdot 0.5724}} = 2.24 > t_{kp} = t_{0.1,10} = 1.81$$

Действительно, есть улучшение!

The En