ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Лекция 15

Проверка статистических гипотез

Статистическая гипотеза — предположение о некоторой закономерности, относящейся к одной или нескольким случайным величинам.

Заключение о правильности или неправильности гипотезы делается по данным выборок на основе того, какова вероятность получить эти выборки (одну или несколько), если гипотеза справедлива.

Гипотеза может состоять из нескольких предположений, часть из которых принимается без доказательства, а другая часть проверяется по данным выборок.

Проверка гипотезы об одинаковости дисперсий

Имеются две выборки из случайных величин x и y, объемом соответственно N_x и N_y .

Их выборочные дисперсии:

$$s_x^2 = \frac{1}{N_x - 1}\sum_{n=1}^{N_x}\left(x_n - \overline{x}\right)^2 = \frac{\sigma_x^2}{N_x - 1}\chi^2_{N_x - 1}$$

$$s_y^2 = \frac{1}{N_y - 1}\sum_{n=1}^{N_y}\left(y_n - \overline{y}\right)^2 = \frac{\sigma_y^2}{N_y - 1}\chi^2_{N_y - 1}$$

Число степеней свободы: $\nu_x = N_x - 1$ и $\nu_y = N_y - 1$

Постулируется: $x \sim \mathbf{N}\left(m^{(1)}, \sigma_x^2\right) \qquad y \sim \mathbf{N}\left(m^{(1)}, \sigma_y^2\right)$

$$\boxed{\text{Гипотеза: } \sigma_x^2 = \sigma_y^2}$$

Критерий Фишера: $\Phi_{\nu_x,\nu_y} = \dfrac{s_x^2}{s_y^2} = \dfrac{\left(N_y - 1\right)\chi^2_{N_x-1}}{\left(N_x - 1\right)\chi^2_{N_y-1}} = K\dfrac{\chi^2_{N_x-1}}{\chi^2_{N_y-1}}$

где $K = \dfrac{\left(N_y - 1\right)}{\left(N_x - 1\right)} = \dfrac{\nu_y}{\nu_x}$

$$w_\chi\left(\chi^2_{N-1}\right) = \frac{1}{2^{\frac{N-1}{2}}\Gamma\left(\frac{N-1}{2}\right)}\left(\chi^2_{N-1}\right)^{\frac{N-1}{2}-1} e^{-\frac{\chi^2_{N-1}}{2}}$$

Имеем: $w_{\chi\chi}\left(\chi^2_{N_x-1},\chi^2_{N_y-1}\right)=w_{\chi}\left(\chi^2_{N_x-1}\right)w_{\chi}\left(\chi^2_{N_y-1}\right)$

Переход к новым переменным: $\chi^2_{N_x-1},\chi^2_{N_y-1}\Rightarrow\Phi_{\nu_x,\nu_y},\chi^2_{N_y-1}$

$$w\left(\Phi_{\nu_x,\nu_y}\right)=\int_0^{\infty}w_{\Phi\chi}\left(\Phi_{\nu_x,\nu_y},\chi^2_{N_y-1}\right)d\left(\chi^2_{N_y-1}\right)$$

После интегрирования получаем распределение Фишера. Плотность распределения в общепринятых обозначениях:

$$w\left(\Phi\,|\,\nu_1,\nu_2\right)=C\frac{\Phi^{\frac{\nu_1}{2}-1}}{\left(\nu_1\Phi+\nu_2\right)^{\frac{\nu_1+\nu_2}{2}}},\text{ где } C=\frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}$$

$$w\left(\Phi \mid \nu_1, \nu_2\right) = Cf\left(\Phi \mid \nu_1, \nu_2\right),$$ где

$$f\left(\Phi \mid \nu_1, \nu_2\right) = \frac{\Phi^{\frac{\nu_1}{2}-1}}{\left(\nu_1 \Phi + \nu_2\right)^{\frac{\nu_1+\nu_2}{2}}}$$

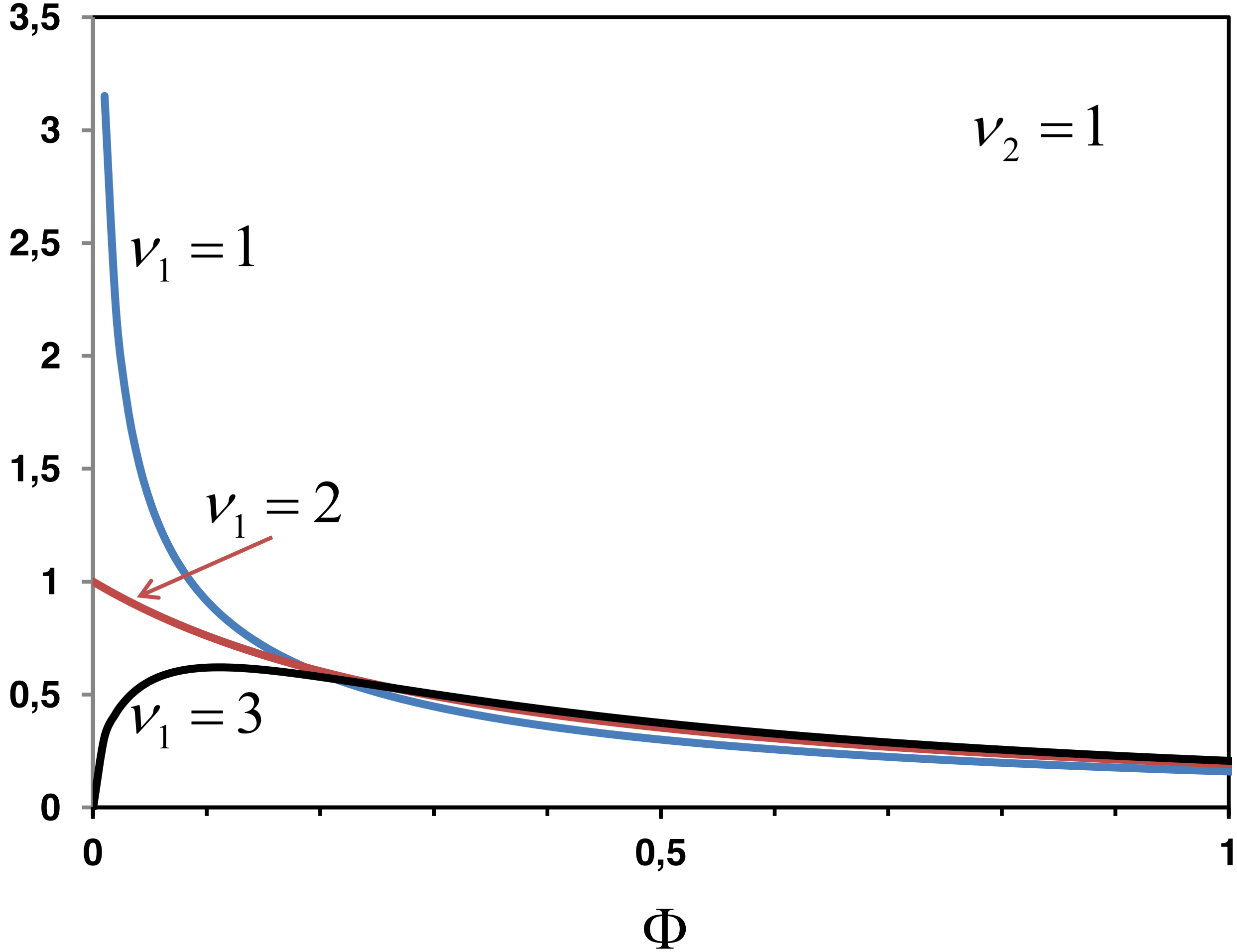
---

1) Вблизи $\Phi = 0$ $\quad f\left(\Phi \mid 1, \nu_2\right) = \dfrac{1}{\Phi^{\frac{1}{2}}\left(\Phi + \nu_2\right)^{\frac{1+\nu_2}{2}}}$

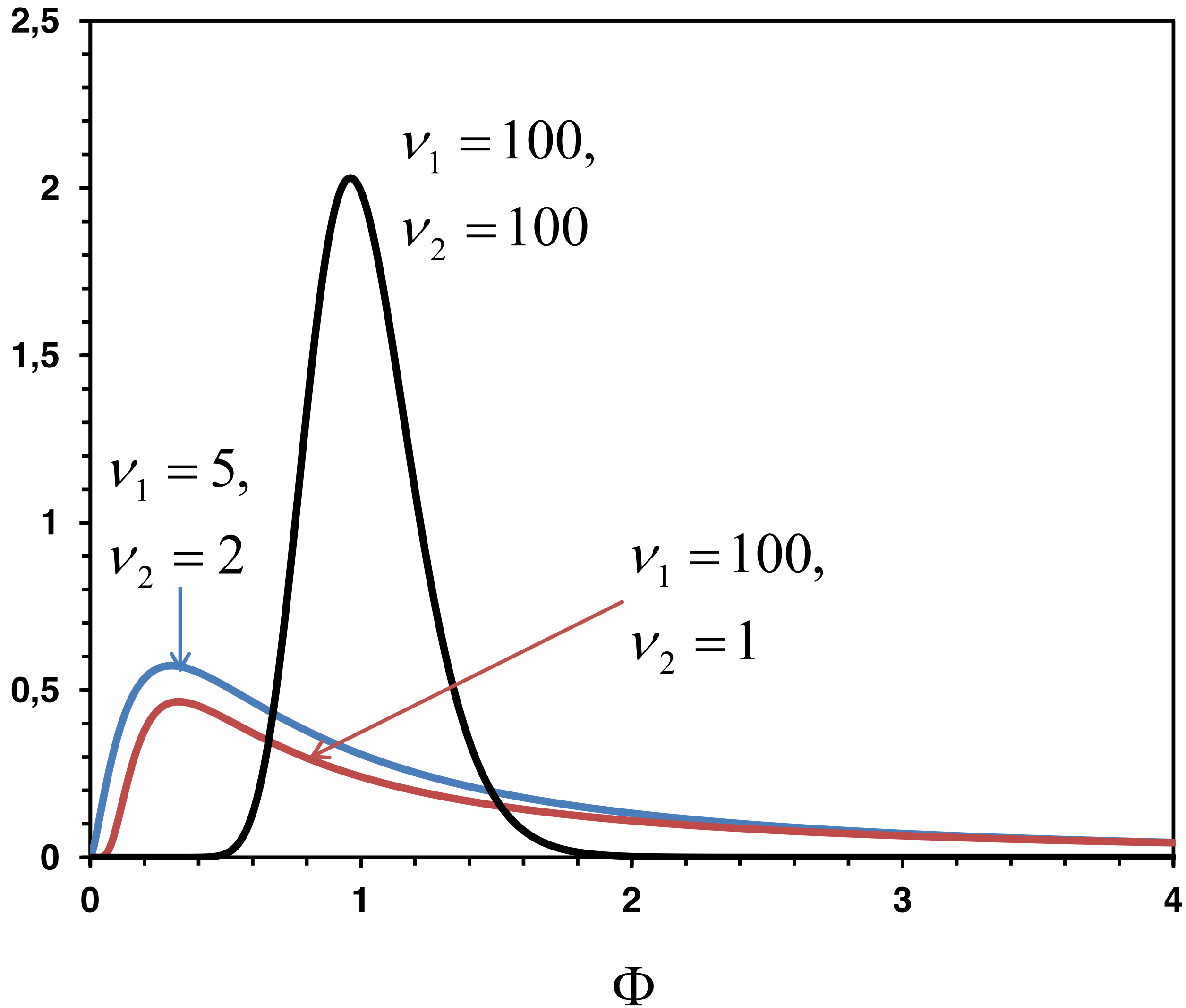
2) $f\left(\Phi \mid 2, \nu_2\right) = \dfrac{1}{\left(2\Phi + \nu_2\right)^{1+\frac{\nu_2}{2}}}$

3) При $\nu_1 > 2$ $\quad f\left(0 \mid \nu_1, \nu_2\right) = 0$

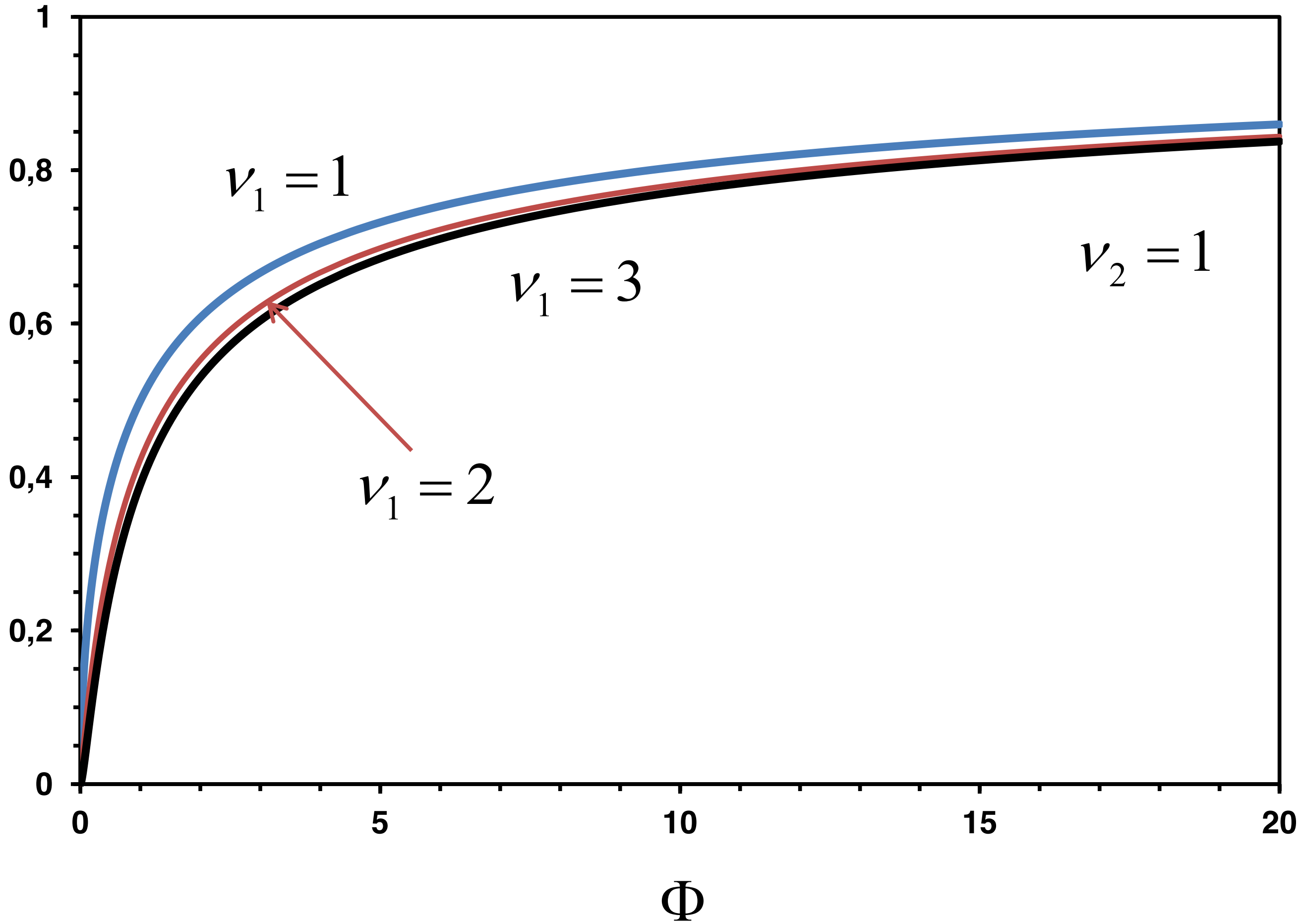
Плотность распределения Фишера



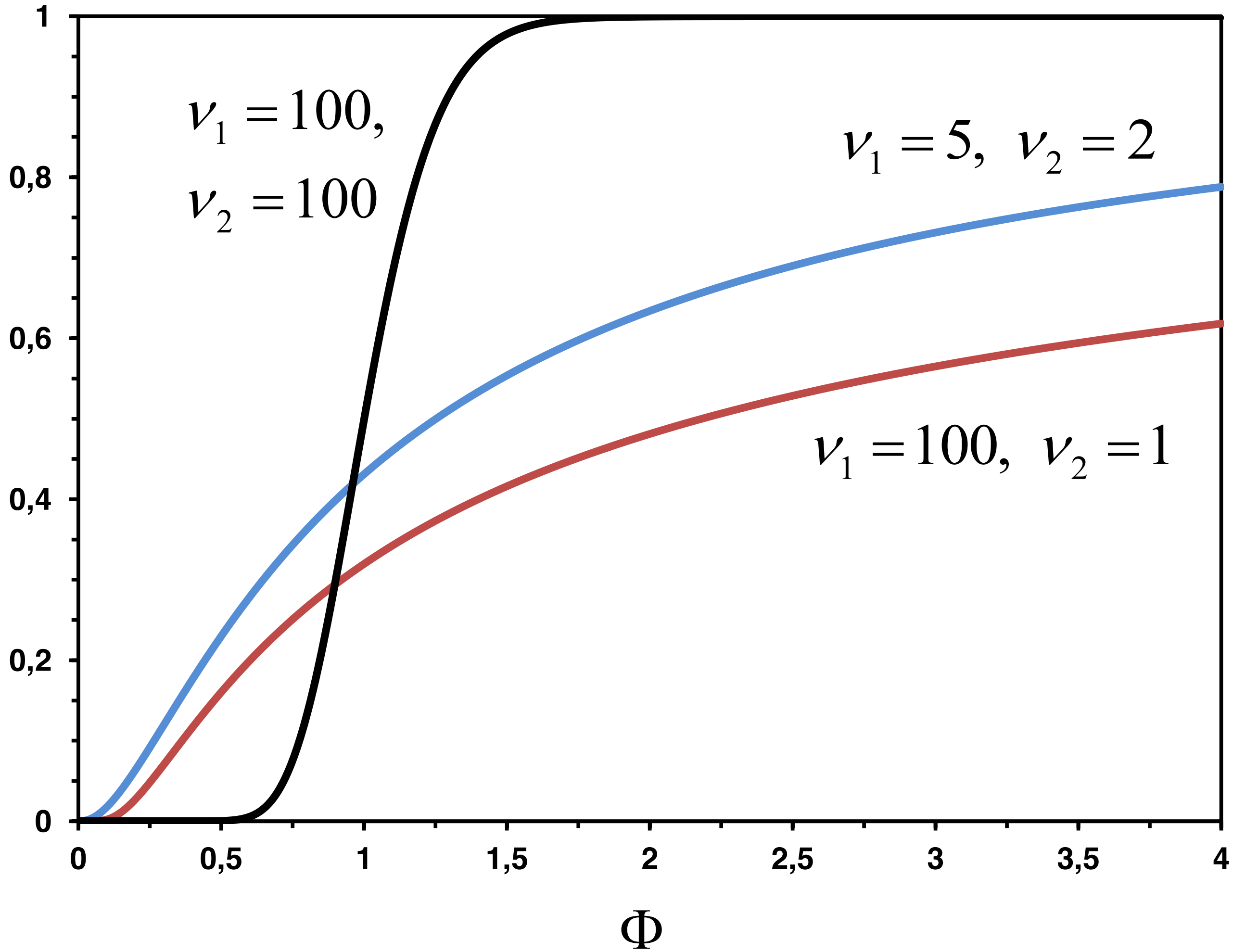
Плотность распределения Фишера



Функция распределения Фишера



Функция распределения Фишера



Пример:

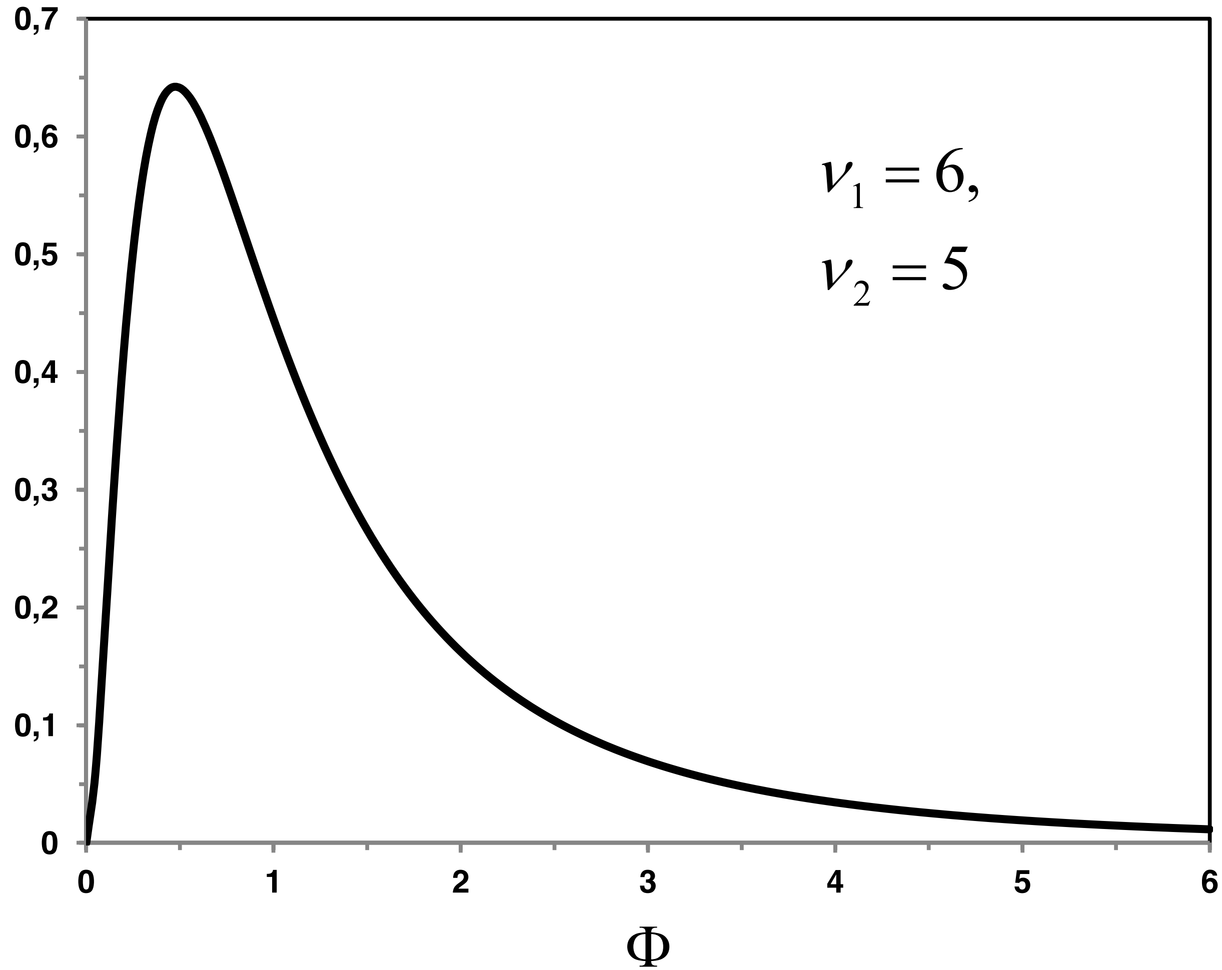
x_1	15.7	10.3	12.6	14.5	12.6	13.8	11.9
x_2	12.3	13.7	10.4	11.4	14.9	12.6	

$$s_1^2 = 3.16, \quad s_2^2 = 2.57$$

$$\nu_1 = 6, \qquad \nu_2 = 5$$

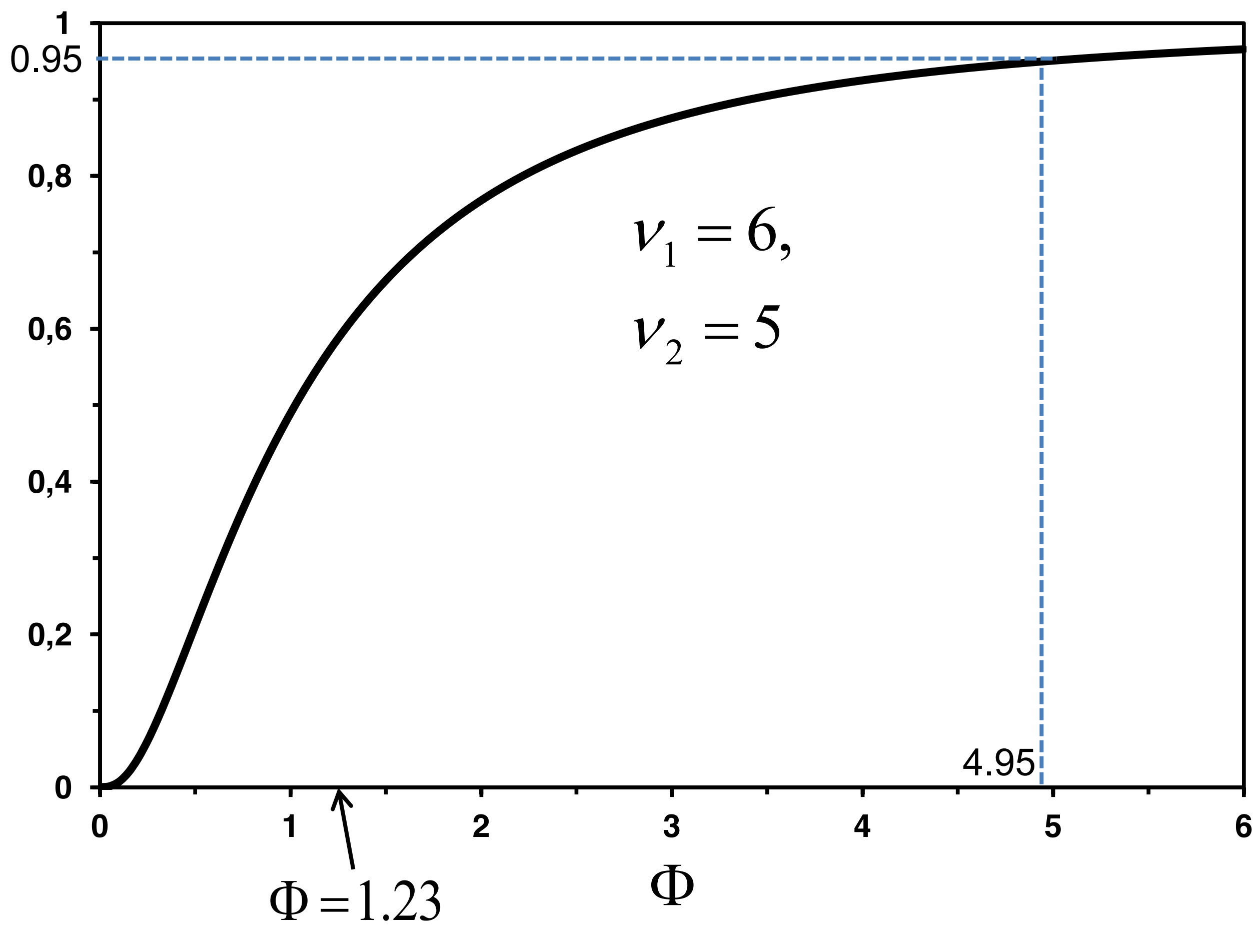
$$\Phi = \frac{s_1^2}{s_2^2} = 1.23$$

Плотность распределения Фишера

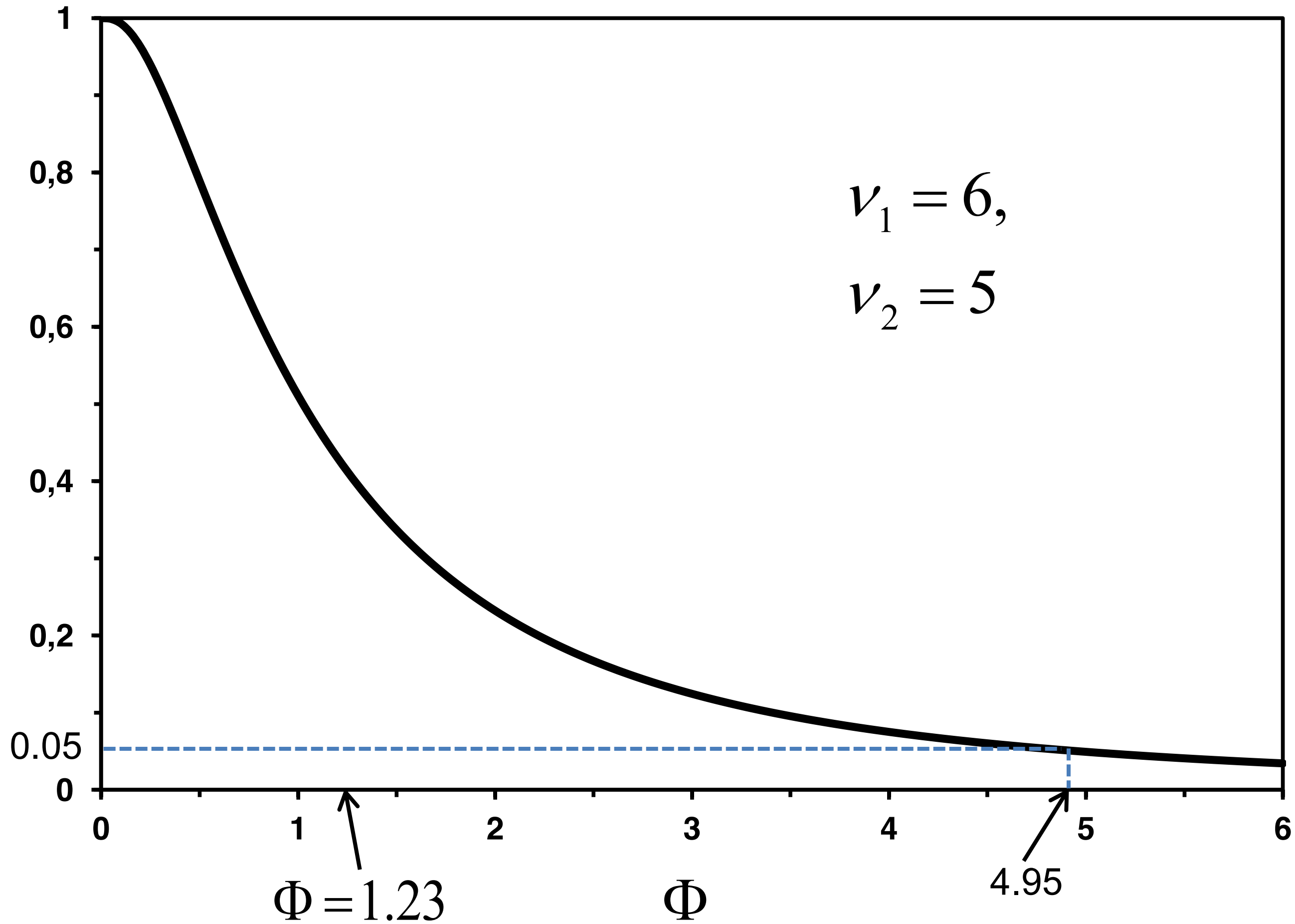


Функция распределения Фишера

$$F\left(\Phi \mid \nu_1, \nu_2\right) = \int_0^{\Phi} w\left(\phi \mid \nu_1, \nu_2\right) d\phi$$



$$F^{\infty}\left(\Phi \mid \nu_1, \nu_2\right) = \int_{\Phi}^{\infty} w\left(\phi \mid \nu_1, \nu_2\right) d\phi$$



Другое рассмотрение тех же данных

x_2	15.7	10.3	12.6	14.5	12.6	13.8	11.9
x_1	12.3	13.7	10.4	11.4	14.9	12.6	

$$s_2^2 = 3.16, \quad s_1^2 = 2.57$$

$$\nu_2 = 6, \qquad \nu_1 = 5$$

$$\Phi = \frac{s_1^2}{s_2^2} = 0.813$$

С вероятностью 0.05 Φ лежит в пределах

$$0 \le \Phi < 0.2$$