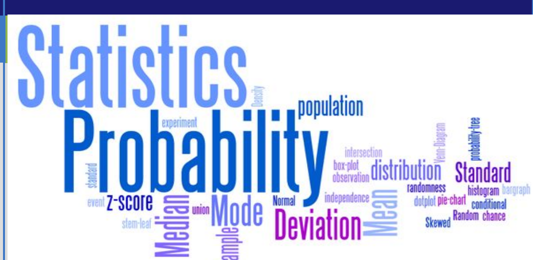
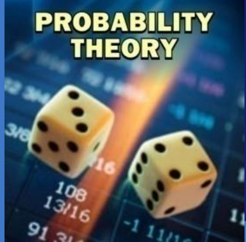
Национальный Университет Узбекистана. Кафедра ТВиМС

ОСНОВНЫЕ РАСПРЕДЕЛЕНИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

нормальное распределение
распределение хи-квадрат
распределение Стьюдента
распределение Фишера



Лекция 6



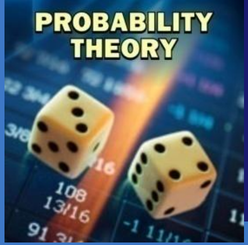
Содержание

1. Нормальное распределение.

2. Хи-квадрат распределение

3. Распределение Стьюдента

4. Распределение Фишера

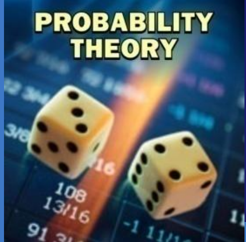


Основные распределения МС

Основные распределения, используемые в математической статистике:

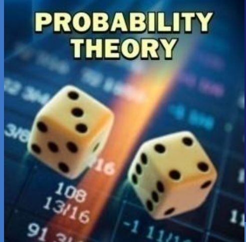
- нормальное распределение,
- распределение хи- квадрат,
- распределения Стьюдента
- распределение Фишера.

При статических исследованиях широко используются случайные величины, имеющие нормальное распределение, распределение χ2 (хи- квадрат), распределения Стьюдента и Фишера.



Нормальное распределение

- **Нормальное распределение**, также называемое **распределением Гаусса**, — распределение вероятностей, которое играет важнейшую роль во многих областях знаний.
- Физическая величина подчиняется нормальному распределению, когда она подвержена влиянию огромного числа случайных помех.
- Ясно, что такая ситуация крайне распространена, поэтому можно сказать, что из всех распределений в природе чаще всего встречается именно нормальное распределение — отсюда и произошло одно из его названий.



Нормальное распределение –N(μ,σ²)

Нормальное распределение, также называемое **распределением Гаусса** или **Гаусса—Лапласа** — распределение вероятностей, которое задаётся функцией плотности вероятности, совпадающей с функцией Гаусса:

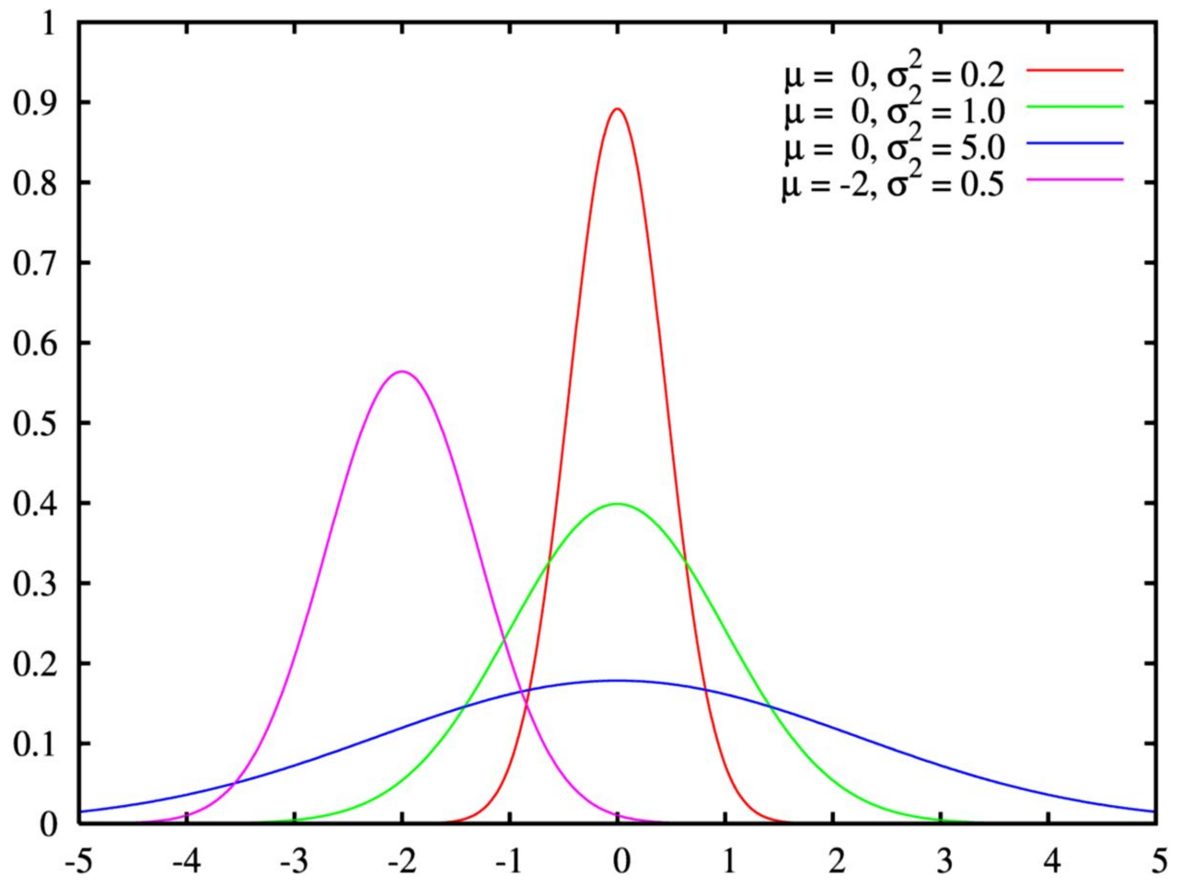
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

где параметр μ — математическое ожидание (среднее значение), медиана и мода распределения, а параметр σ— среднеквадратическое отклонение (σ^2 —дисперсия) распределения.

Плотность распределения

Плотность нормального распределения:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$



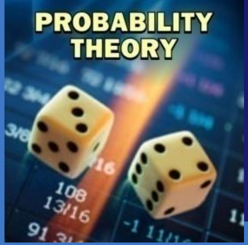
Нормальное распределение зависит от двух параметров —**(μ)** ***смещения* и (σ)** ***масштаба***, то есть является с математической точки зрения не одним распределением, а целым их семейством.

Плотность распределения

Графики плотностей нормальных распределений с нулевым средним и разными отклонениями (σ=0.5, σ=1, σ =2).

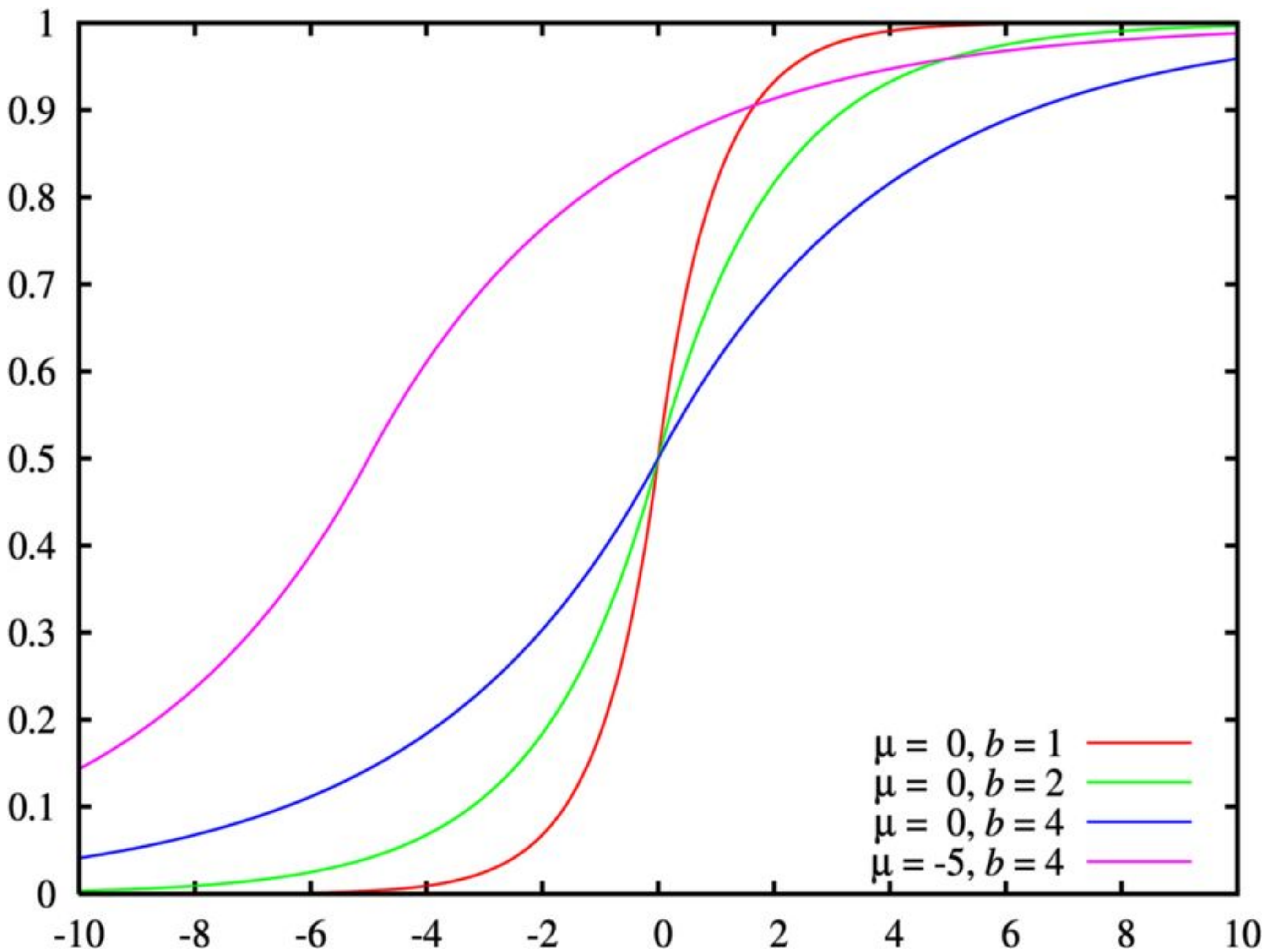
Плотность распределения

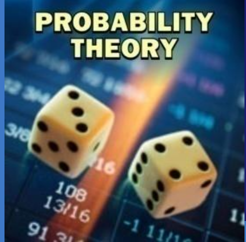
Графики плотностей двух нормальных распределений N(-2,2) и N(3,2).



Функция распределения

$$\mathbf{F}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$



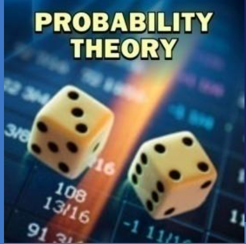


Нормальное распределение

Параметры	μ -коэффициент сдвига $\sigma > 0$ - коэффициент масштаба
Носитель	$x \in (-\infty; +\infty)$
Плотность вероятности	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Функция распределения	$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$
Характеристическая функция	$\phi_X(t) = \exp\left(\mu i t - \frac{\sigma^2 t^2}{2}\right)$

Характеристики нормального распределения

Математическое ожидание	μ
Медиана	μ
Мода	μ
Среднеквадратическое отклонение	σ
Дисперсия	σ^2
Коэффициент асимметрии	0
Коэффициент эксцесса	0



Моменты

- Если X имеет нормальное распределение, то для неё существуют (конечные) моменты при всех p с действительной частью больше −1.

$$E\left[(X-\mu)^p\right]=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-\mu)^p\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$$

- Для неотрицательных целых p, центральные моменты таковы:

$$E\left[(X-\mu)^p\right]=\begin{cases}0, & p=2n+1;\\ \sigma^p(p-1)!!, & p=2n.\end{cases}$$

- Центральные абсолютные моменты для неотрицательных целых p таковы:

$$E\left[|X-\mu|^p\right]=\sigma^p(p-1)!!\cdot\begin{cases}\sqrt{\frac{2}{\pi}}, & p=2n+1;\\ 1, & p=2n.\end{cases}=\sigma^p\frac{2^{\frac{p}{2}}\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}$$

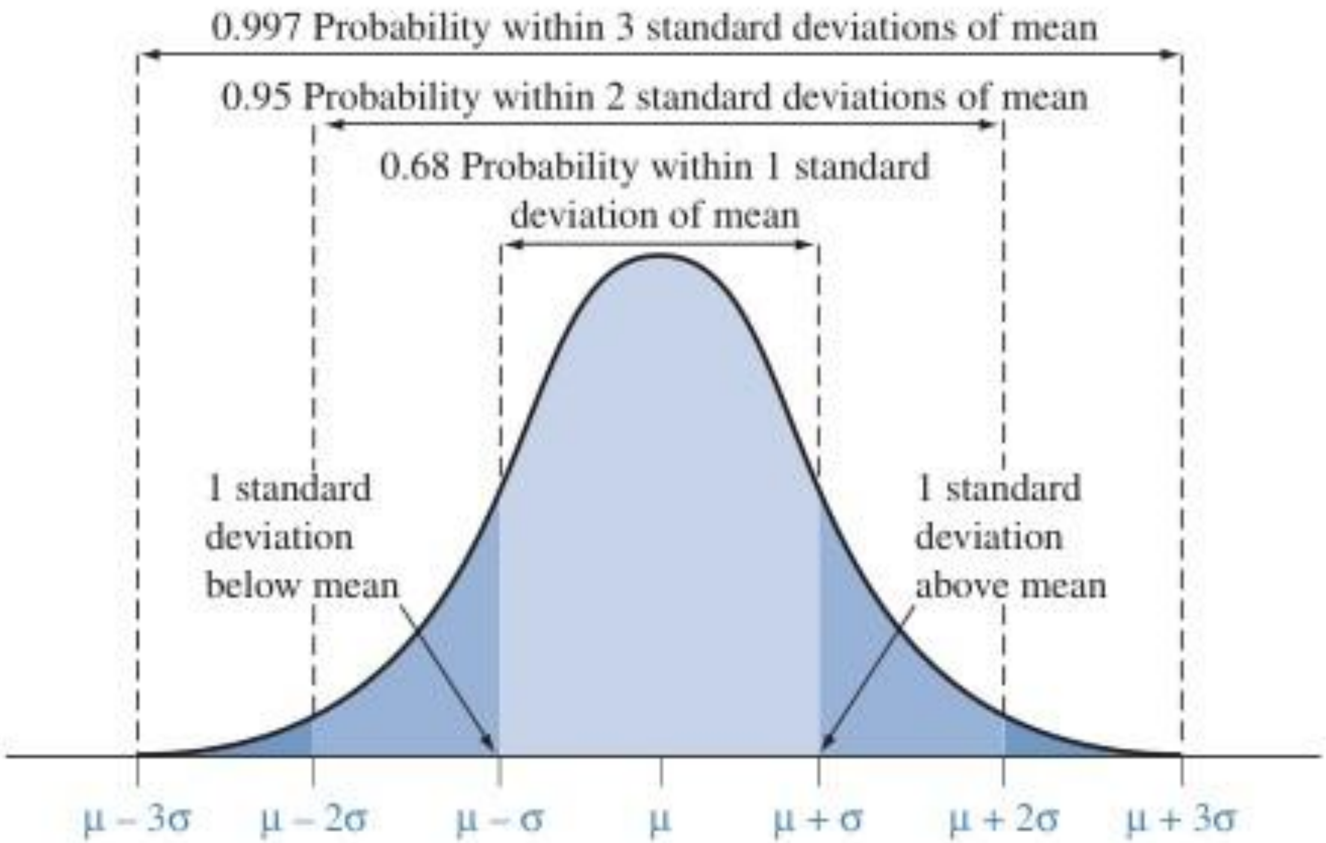
Правило сигм

- Нормально распределенная случайная величина с большой вероятностью принимает значения, близкие к своему математическому ожиданию, что выражается правилом сигм:

$$P(|X-m|<k\sigma)=\begin{cases}0{,}6827, & k=1\\ 0{,}9545, & k=2\\ 0{,}9973, & k=3\end{cases}$$

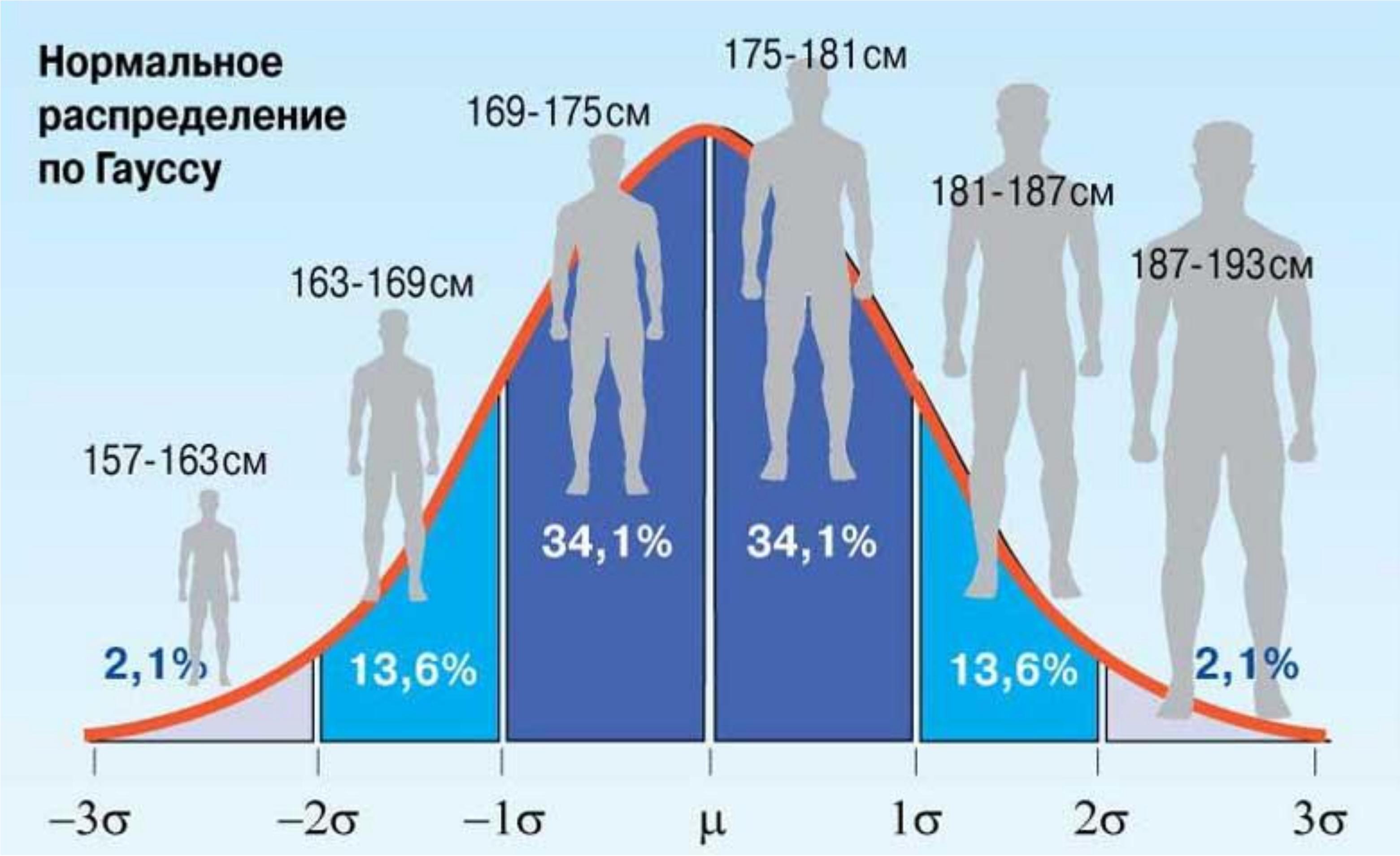
- Чаще всего используется правило трех сигм, т.е. k = 3.

Правило сигм



Для нормального распределения значения, отличающиеся от среднего на число, меньшее чем одно стандартное отклонение, составляют 68,27 % популяции.
В то же время значения, отличающиеся от среднего на два стандартных отклонения, составляют 95,45 %, а на три стандартных отклонения — 99,73 %.

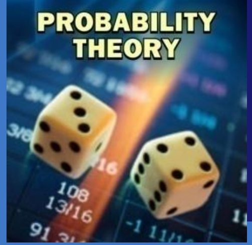
Правило сигм



Стандартное нормальное распределение - N(0,1)

- **Стандартным нормальным распределением** называется нормальное распределение с математическим ожиданием 0 и стандартным отклонением 1.

Обозначение	N(0,1)
Параметры	0 — коэффициент сдвига 1 — коэффициент масштаба
Носитель	$x \in (-\infty; +\infty)$
Плотность вероятности	$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$
Функция распределения	$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$
Математическое ожидание, медиана, мода	0
Дисперсия, среднее квадратическое отклонение	1
Коэффициент асимметрии, Коэффициент эксцесса	0



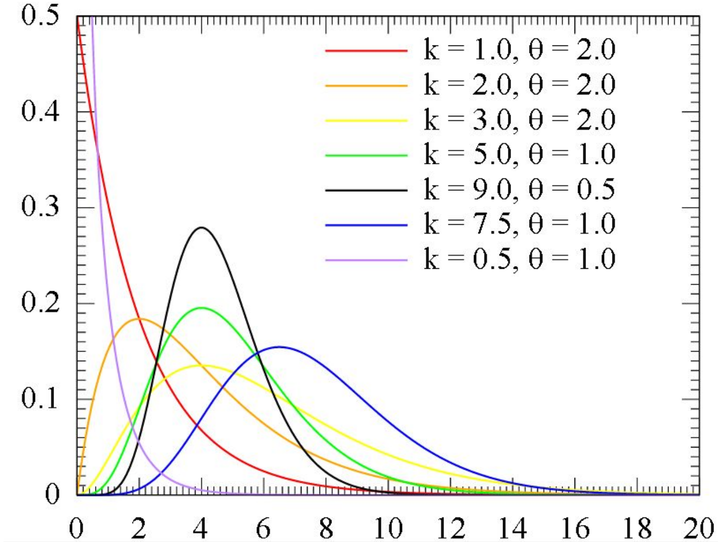
Гамма распределение

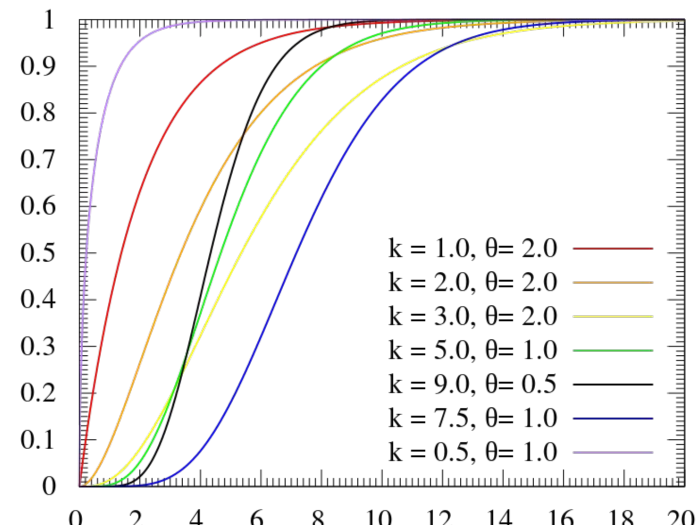
Пусть распределение случайной величины X задаётся плотностью вероятности, имеющей вид

$$f_X(x) = \begin{cases} x^{k-1} \dfrac{e^{-x/\theta}}{\theta^k \, \Gamma(k)}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

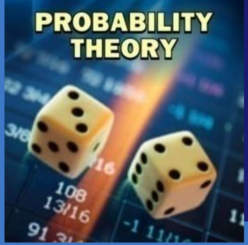
, где $\Gamma(k)$ — гамма-функция Эйлера.

Тогда говорят, что случайная величина X имеет гамма-распределение с положительными параметрами θ и k. Пишут $X \sim \Gamma(k, \theta)$.





Функция распределения $\dfrac{1}{\Gamma(k)}\gamma\left(k, \dfrac{x}{\theta}\right)$



Хи-квадрат распределение

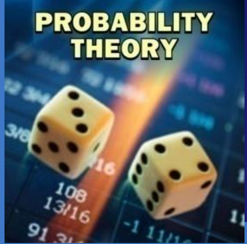
Пусть $\xi_1,\dots,\xi_k$ — совместно независимые стандартные нормальные случайные величины, то есть: $\xi_i \sim N(0,1)$.

Тогда случайная величина

$$\chi^2 = \xi_1^2 + \xi_2^2 + \dots + \xi_k^2$$

имеет **распределение хи-квадрат** с k степенями свободы, то есть

$$\chi^2 = \sum_{i=1}^{k} \xi_i^2 \sim \chi^2(k)$$



Распределение хи-квадрат

Распределение хи-квадрат является частным случаем гамма распределения, и его плотность имеет вид:

$$f_{\chi^2(k)}(x) \equiv \Gamma\left(\frac{k}{2}, 2\right) = \frac{(1/2)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},$$

где Г(k/2,2) означает гамма-распределение, а Г (k/2) — гамма функции.

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx.$$

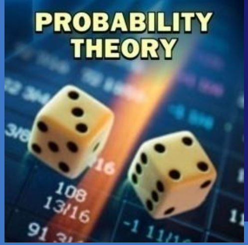
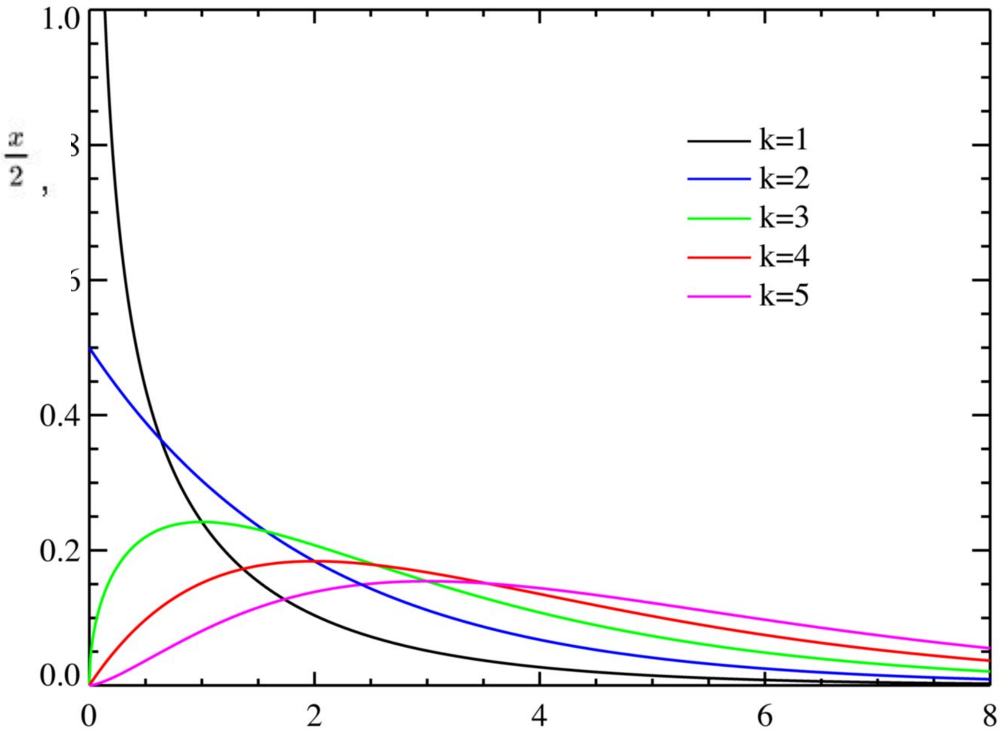
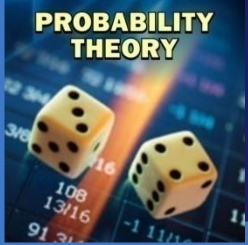


График плотности распределения

$$f_{\chi^2(k)}(x) \equiv \Gamma\left(\frac{k}{2}, 2\right) = \frac{(1/2)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},$$



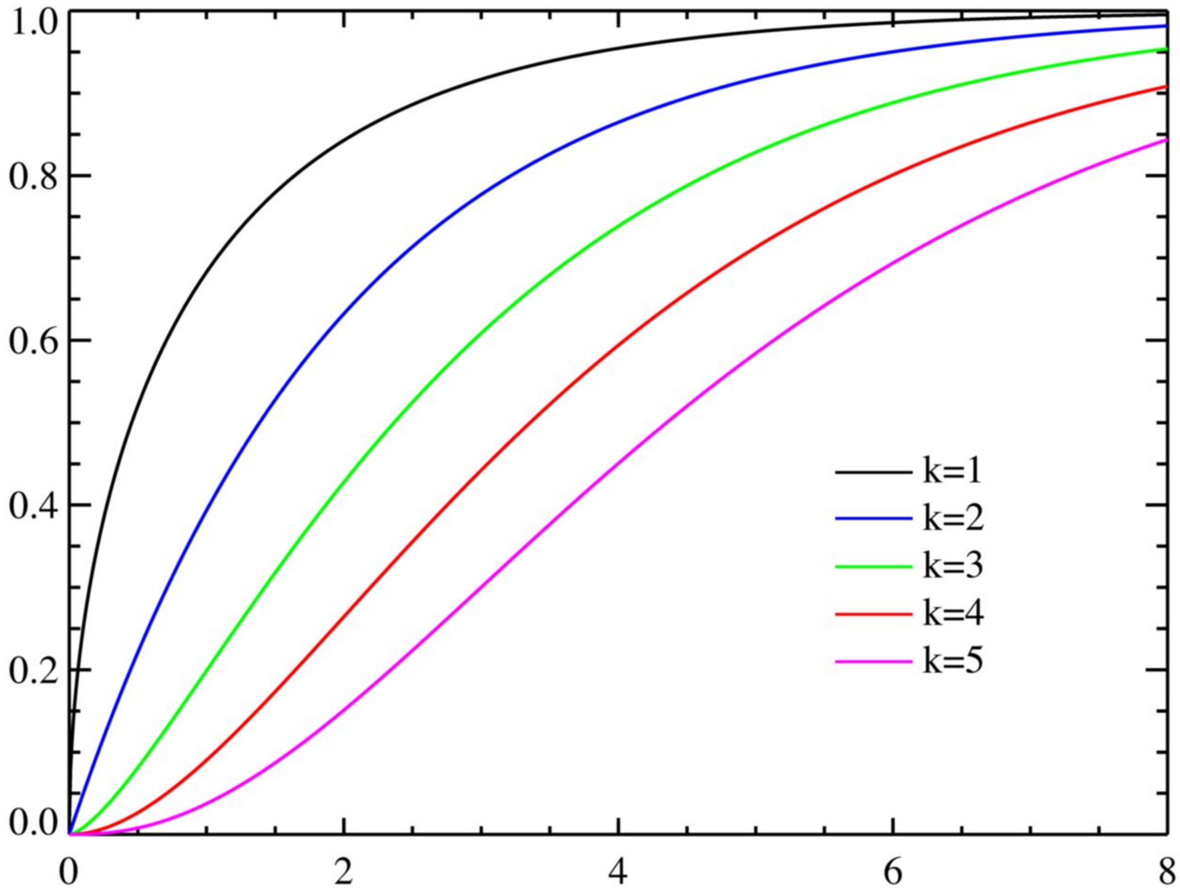
Графики плотности распределения χ2 с k степенями свободы асимметричны и, начиная с k=2, имеют по одному максимуму в точке x=k−2. Причем с ростом k кривая плотности приближается к симметричной функции.

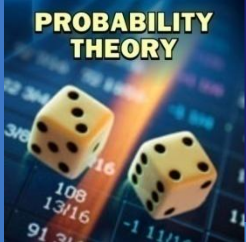


Функция распределения

Функция распределения $\dfrac{\gamma(k/2, x/2)}{\Gamma(k/2)}$

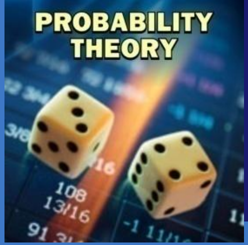
$$\gamma(a,z) = \int_0^z e^{-t} t^{a-1}\, dt.$$





Характеристики распределения

Обозначение	$\chi^2(k)$ или χ^2_k
Параметры	$k > 0$ — число степеней свободы
Носитель	$x \in [0; +\infty)$
Плотность вероятности	$\dfrac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}$
Функция распределения	$\dfrac{\gamma(k/2, x/2)}{\Gamma(k/2)}$
Математическое ожидание	k
Медиана	примерно $k - 2/3$
Мода	$k - 2$ если $k \geq 2$
Дисперсия	$2\,k$
Коэффициент асимметрии	$\sqrt{8/k}$
Коэффициент эксцесса	$12/k$



Распределение Стьюдента

- **Распределение Стьюдента** (**t-распределение**) в теории вероятностей — это однопараметрическое семейство абсолютно непрерывных распределений.
- Уильям Сили Госсет первым опубликовал работы, посвящённые этому распределению, под псевдонимом «Стьюдент».

Распределение Стьюдента

Пусть $Y_0, Y_1, \ldots, Y_n$ — независимые стандартные нормальные случайные величины, такие что $Y_i \sim N(0,1)$, $i=0,\ldots,n$. Тогда распределение случайной величины t, где

$$t = \frac{Y_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2}}$$

называется **распределением Стьюдента с n степенями свободы $t \sim t(n)$.**

Распределение Стьюдента

Пусть случайная величина Y имеет распределение N(0,1), а независимая от Y случайная величина Z принадлежит $\chi^2(n)$ (имеет распределение χ^2 с n степенями свободы).

Тогда случайная величина

$$t=\frac{Y}{\sqrt{\frac{Z}{n}}}$$

имеет распределение Стьюдента с n степенями свободы (t(n) - распределение)

Плотность распределения

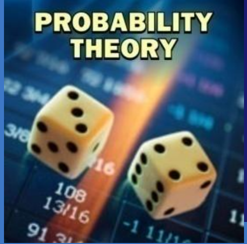
$$f_t(y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{y^2}{n}\right)^{-\frac{n+1}{2}},$$

где Γ — гамма-функция Эйлера. Таким образом:

$$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})} = \frac{(n-1)(n-3)\cdots 5\cdot 3}{2\sqrt{n}(n-2)(n-4)\cdots 4\cdot 2},$$ для чётных n

и соответственно

$$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})} = \frac{(n-1)(n-3)\cdots 4\cdot 2}{\pi\sqrt{n}(n-2)(n-4)\cdots 5\cdot 3},$$ для нечётных n.



Плотность распределения

Также плотность распределения Стьюдента можно выразить воспользовавшись бета-функцией Эйлера B:

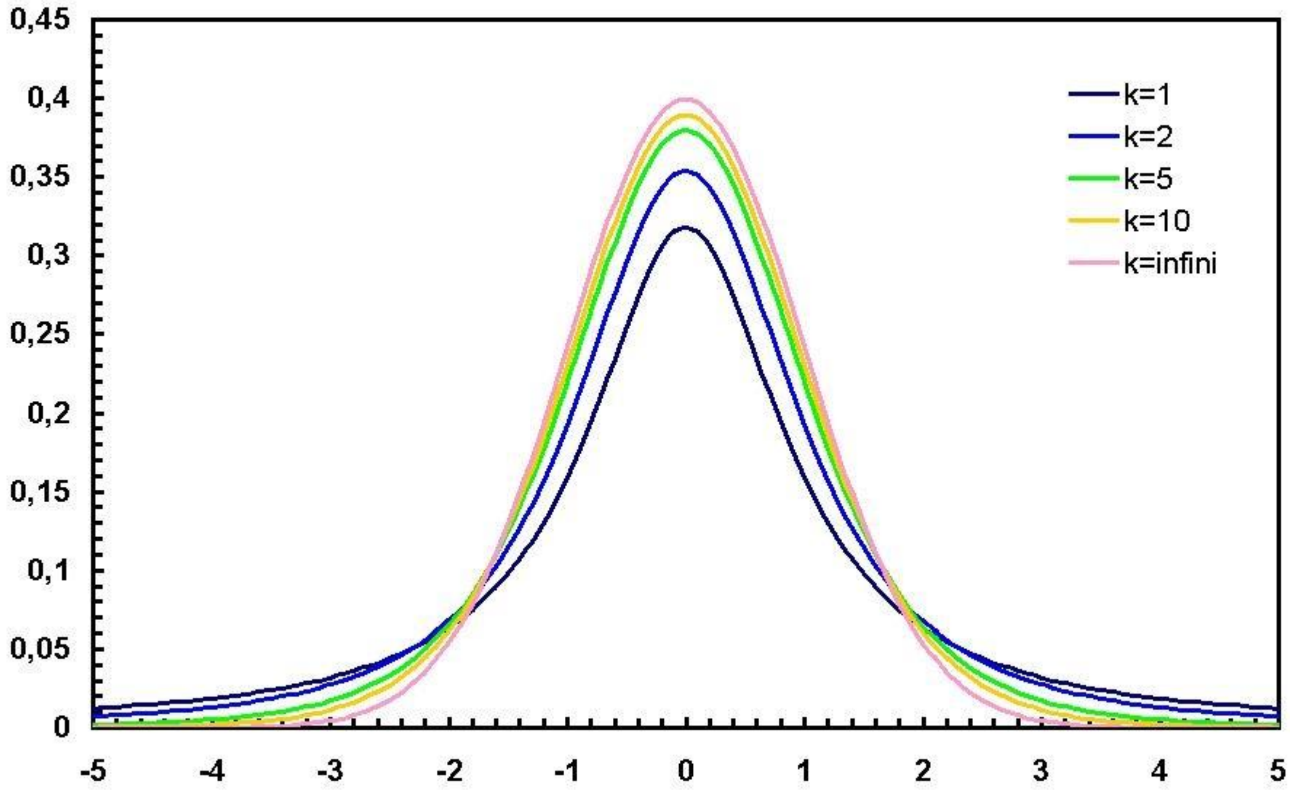
$$f_t(y) = \frac{1}{\sqrt{n}\,\mathrm{B}\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{y^2}{n}\right)^{-\frac{n+1}{2}}.$$

Бета-функцией (B-функцией, бета-функцией Эйлера или интегралом Эйлера I рода) называется следующая функция от двух переменных:

$$\mathrm{B}(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$$

График плотности распределения

С ростом n распределение Стьюдента приближается к N(0,1) .



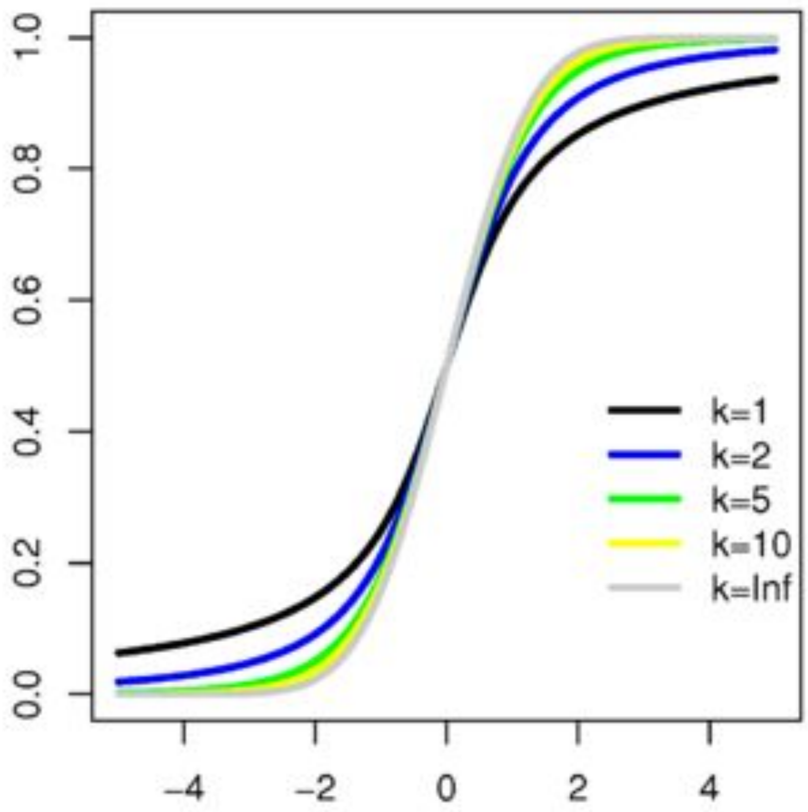
Графики плотности случайной величины, имеющей распределение Стьюдента, при любом n = 1,2,.... симметричны относительно оси ординат, поэтому при любом n =1,2,..... математическое ожидание равно нулю.

Функция распределения

Функция распределения

$$\frac{1}{2} + x\Gamma\left(\frac{n+1}{2}\right) \times \frac{{}_2F_1\left(\frac{1}{2}, \frac{n+1}{2}; \frac{3}{2}; -\frac{x^2}{n}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}$$

где ${}_2F_1$ — гипергеометрическая функция



Характеристики распределения

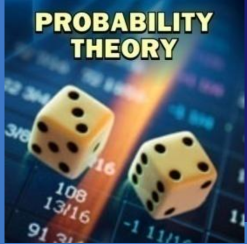
Обозначение	$\mathrm{t}(n)$
Параметры	$n > 0$ — число степеней свободы
Носитель	$x \in (-\infty; +\infty)$
Плотность вероятности	$\dfrac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})\,(1+\frac{x^2}{n})^{\frac{n+1}{2}}}$
Функция распределения	$\dfrac{1}{2} + x\Gamma\left(\dfrac{n+1}{2}\right) \times \dfrac{{}_2F_1\left(\frac{1}{2}, \frac{n+1}{2}; \frac{3}{2}; -\frac{x^2}{n}\right)}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})}$, где ${}_2F_1$ — гипергеометрическая функция

Математическое ожидание	0, если $n > 1$
Медиана	0
Мода	0
Дисперсия	$\dfrac{n}{n-2}$, если $n > 2$
Коэффициент асимметрии	0, если $n > 3$
Коэффициент эксцесса	$\dfrac{6}{n-4}$, если $n > 4$

Распределение Фишера

Пусть Y_1, Y_2 — две независимые случайные величины, имеющие распределение хи-квадрат: $Y_i \sim \chi^2(d_i)$, где $d_i \in \mathbb{N},\ i = 1, 2$. Тогда распределение случайной величины

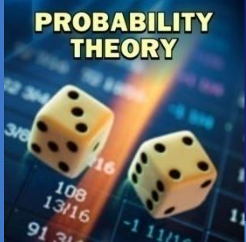
$F = \dfrac{Y_1/d_1}{Y_2/d_2}$ называется распределением Фишера (распределением Снедекора) со степенями свободы d_1 и d_2. Пишут $F \sim \mathrm{F}(d_1, d_2)$.



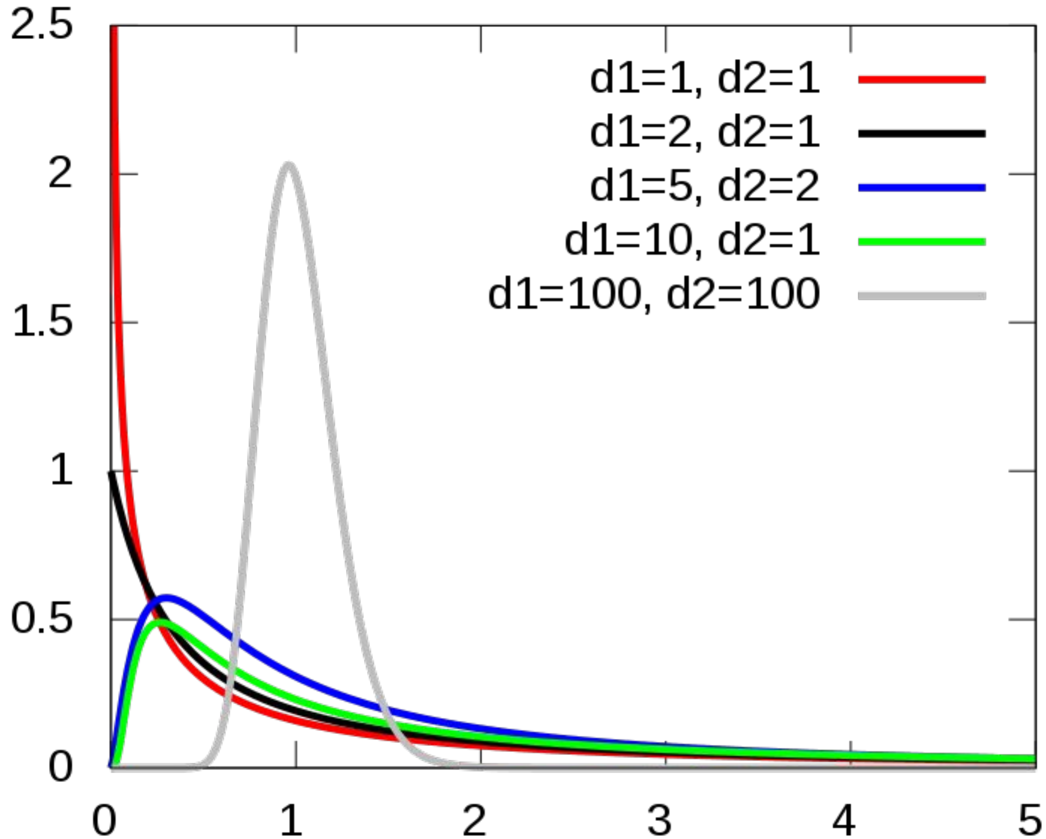
Плотность распределения

Плотностьраспределения X~F(n1;n2)

$$f(x)=\begin{cases}\dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)\cdot n_1^{\frac{n_1}{2}}\cdot n_2^{\frac{n_2}{2}}}{\Gamma\left(\dfrac{n_1}{2}\right)\cdot\Gamma\left(\dfrac{n_2}{2}\right)}\cdot x^{\frac{n_1}{2}-1}\cdot\left(n_2+n_1x\right)^{-\frac{n_1+n_2}{2}}, x>0\\ 0, x\le 0\end{cases}.$$

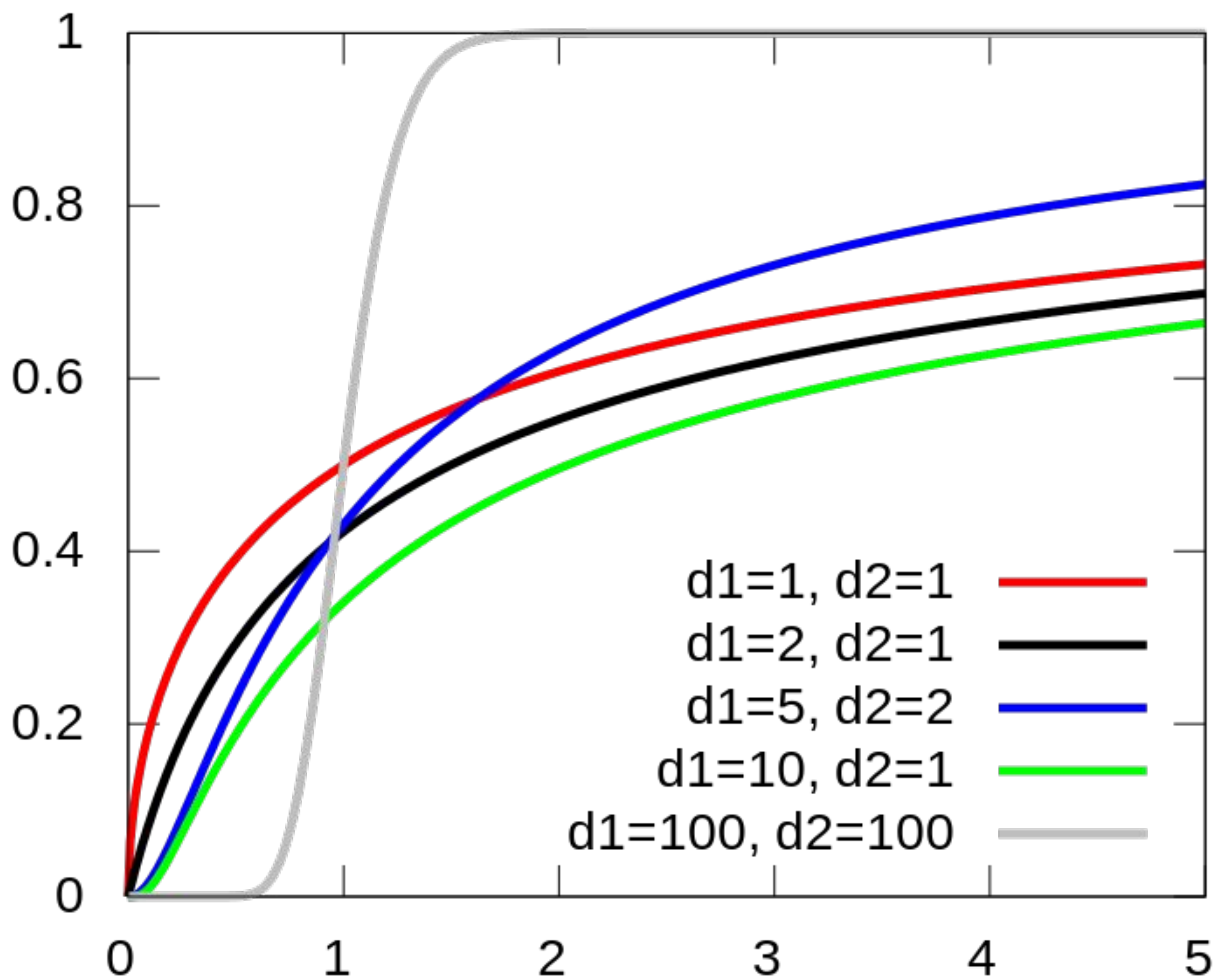


Плотность распределения



Графики плотности распределения случайной величины асимметричны, имеют длинные "хвосты" и достигают максимума вблизи точки x=1.

Функция распределения



PROBABILITY THEORY

Характеристики распределения

Математическое ожидание	$\frac{d_2}{d_2-2}$, если $d_2>2$
Мода	$\frac{d_1-2}{d_1}\;\frac{d_2}{d_2+2}$, если $d_1>2$
Дисперсия	$\frac{2\,d_2^2\,(d_1+d_2-2)}{d_1\,(d_2-2)^2\,(d_2-4)}$, если $d_2>4$
Коэффициент асимметрии	$\frac{(2d_1+d_2-2)\sqrt{8(d_2-4)}}{(d_2-6)\sqrt{d_1\,(d_1+d_2-2)}}$, если $d_2>6$

Спасибо за внимание !