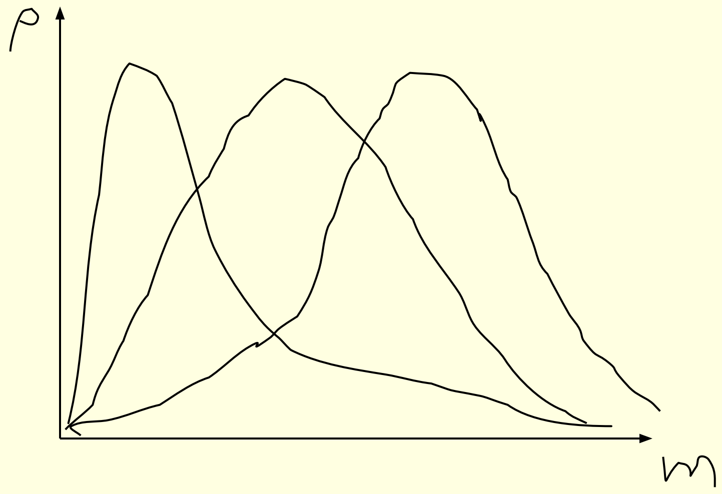
ЛЕКЦИЯ 4



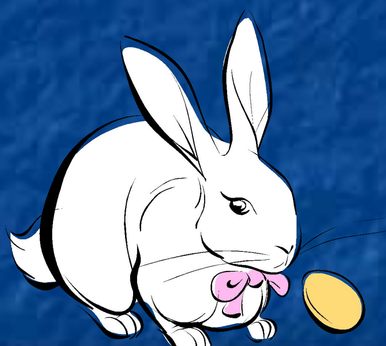
ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ СЛУЧАЙНЫХ ВЕЛИЧИН

4.1. Эмпирические и теоретические распределения вероятностей случайных величин

Пример вариационного ряда: данные о плодовитости кроликов

Количество крольчат (x_i): 2 3 4 5

Частота варианты (f_i): 1 2 5 2



Почему важен анализ теоретических распределений вероятностей???

- Предполагая, какие факторы влияют на исследуемое явление, можно также предполагать, как будут распределяться *экспериментальные данные*;
- Если же получаемые данные не соответствуют ожидаемому распределению, следует заключить, что предполагавшиеся факторы *не оказывают влияние на данное явление*.

4.2. Вероятности и их свойства

Понятия теории вероятностей:

Событие – результат (=исход) отдельного испытания.

Понятия теории вероятностей:

Несколько событий называются <u>несовместимыми</u>, если в условиях испытания каждый раз возможно наступление только одного из них. Иначе события будут <u>совместимыми</u>.

Понятия теории вероятностей:

Два события называются <u>противоположными</u>, если наступление любого из них исключает появление другого

Понятия теории вероятностей:

- **Достоверное событие** – происходит неизбежно при каждом испытании;
- **Невозможное событие** – в заданных условиях произойти не может;
- **Случайное событие** – может произойти, но может и не произойти в данных условиях.

Понятия теории вероятностей:

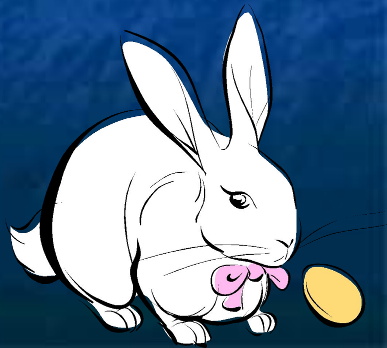
<u>Вероятностью</u> называется отношение числа случаев, благоприятствующих наступлению ожидаемого события, к числу всех возможных исходов:

$$P(A) = m/n,$$

Пример:

В корзине сидят 5 белых и 10 черных кроликов. Какова вероятность вытащить наугад белого кролика?

$$P = 5/15 = 0,33$$



Границы возможных значений вероятностей:

$$0 \le P \le 1$$

p – вероятность ожидаемого события;

q – вероятность противоположного ему события;

$$\Rightarrow \boxed{p + q = 1}$$

Свойства вероятностей:

1. **Вероятность наступления одного из двух (все равно какого) или нескольких независимых и несовместных событий А, В, С … К равна сумме их вероятностей:**

Р(А+В+С+…+К) = Р(А) + Р(В) + Р(С) + … + Р(К)

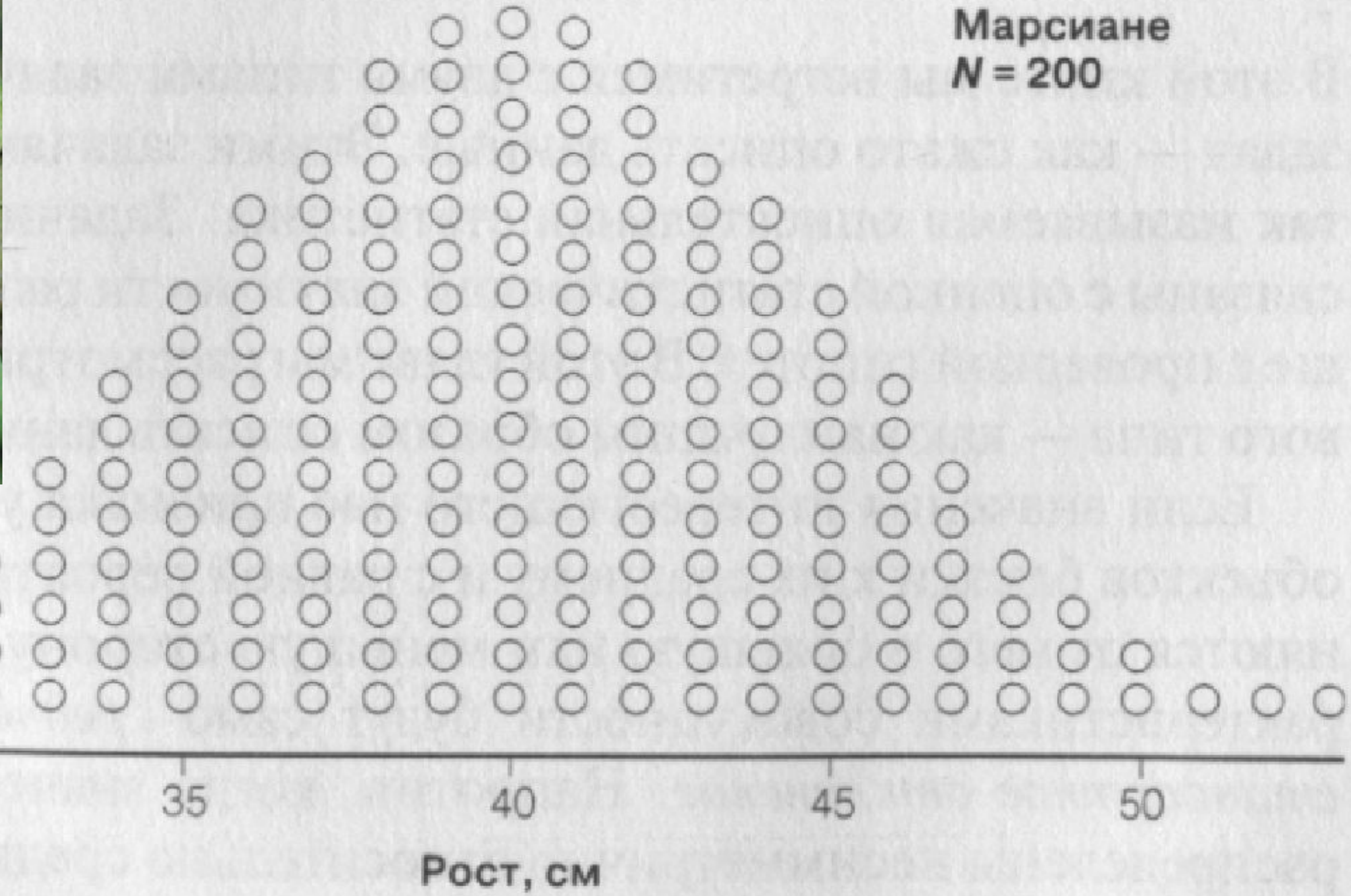
Свойства вероятностей:

2. Вероятность совместного появления двух или нескольких независимых событий равна произведению вероятностей этих событий:

P(A, B, C, …, K) = P(A) P(B) P(C) … P(K)

4.3. Закон нормального распределения вероятностей

Распределение марсиан по росту



Нормальное распределение формируется тогда, когда некая величина отклоняется от среднего под действием множества слабых, независимых друг от друга факторов

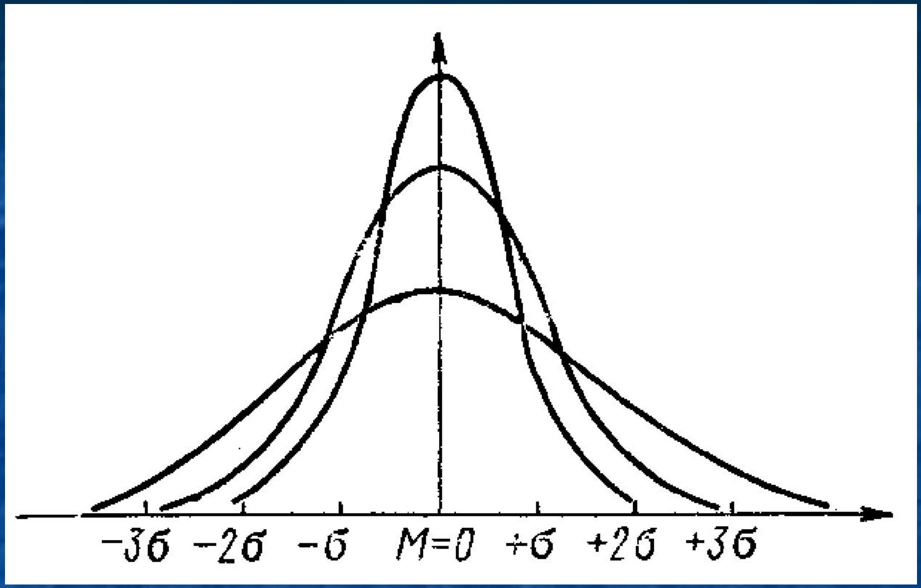
Математическое описание закона нормального распределения вероятностей:

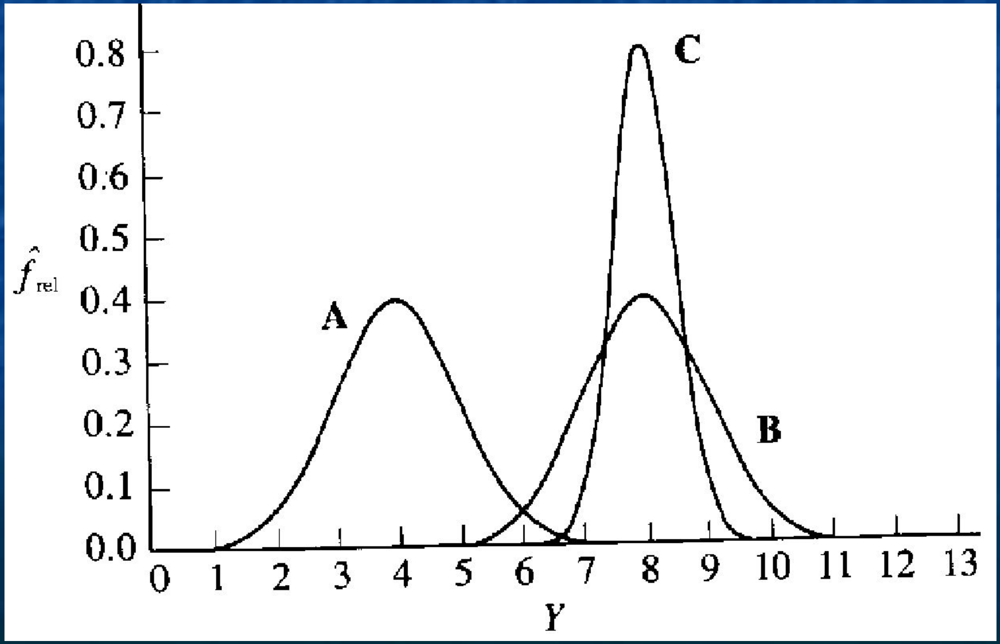
$$P(x_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2[(x_i - \mu)/\sigma]^2}$$

μ **- генеральная средняя (=математическое ожидание);**

σ **- стандартное отклонение.**

Нормальные кривые:





Коэффициент асимметрии нормальной кривой (*skewness*):

$$g_1 = \frac{\sum (x_i - \bar{x})^3}{ns^3}$$

Коэффициент эксцесса нормальной кривой (*kurtosis*):

$$g_2 = \frac{\sum (x_i - \bar{x})^4}{ns^4} - 3$$

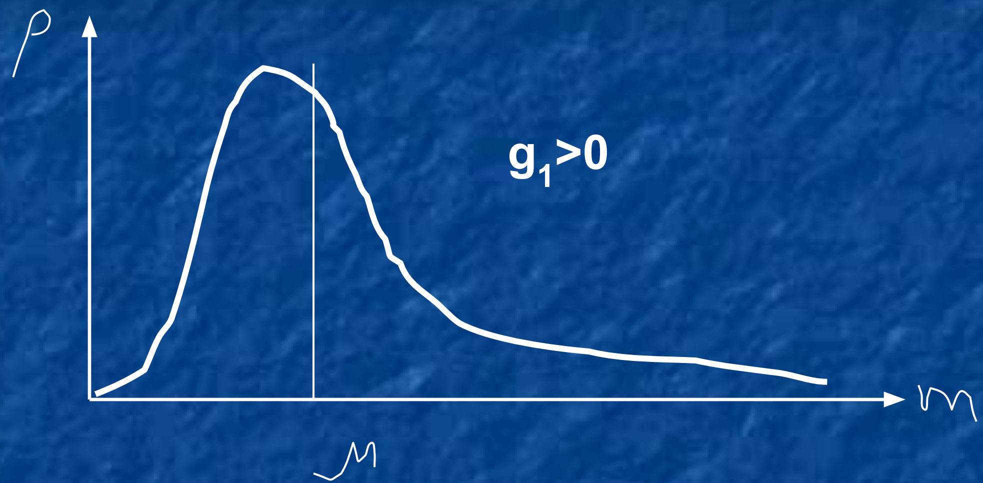
Интерпретация коэффициентов асимметрии и эксцесса:

При g_1 и $g_2 = 0$ кривая строго симметрична;

При $g_1 < 0$ кривая скошена влево (и наоборот);

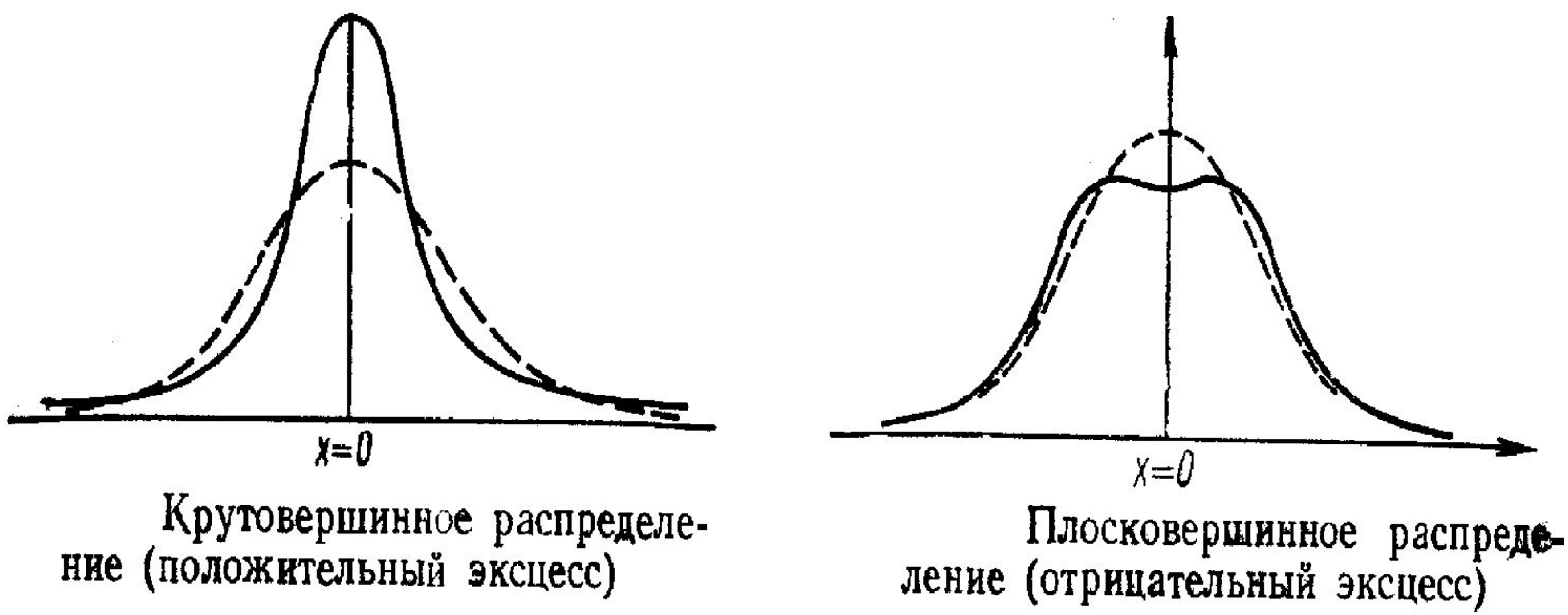
При $g_2 < 0$ кривая обладает плоской вершиной (и наоборот – «бока» кривой крутые).

Нормальная кривая, скошенная вправо



Нормальная кривая, скошенная влево

Круто- и плосковершинные нормальные кривые:

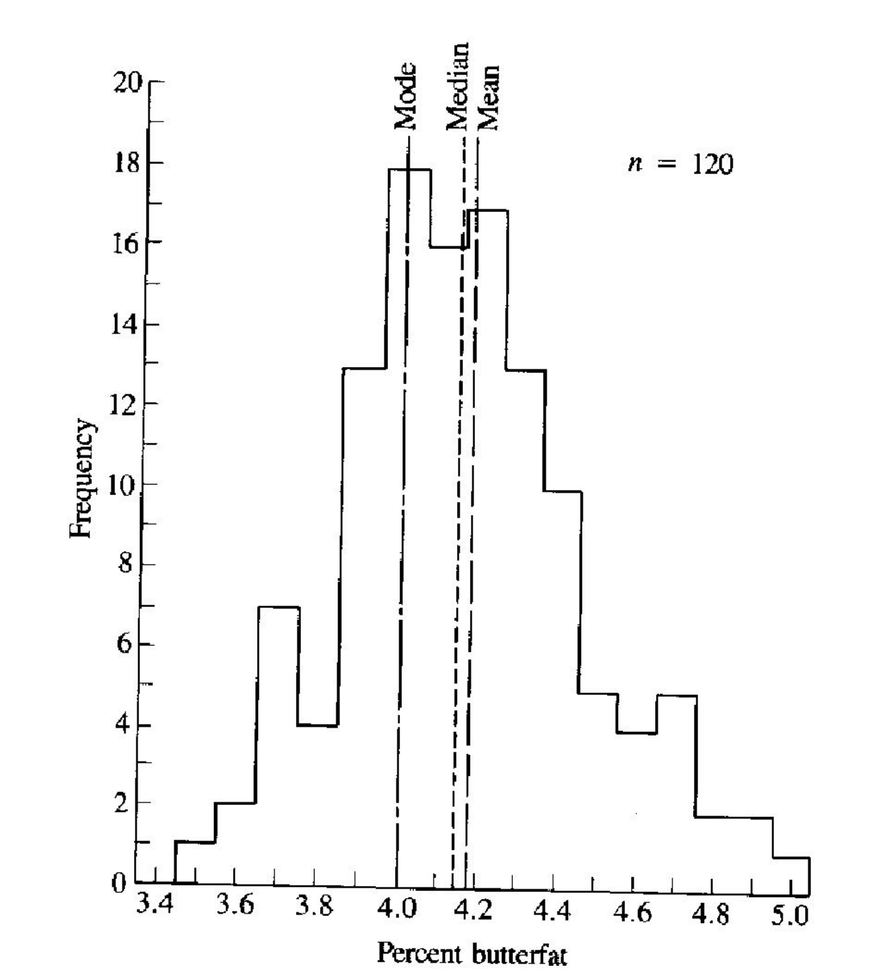


$g_2>0$

$g_2<0$

Свойства нормального распределения:

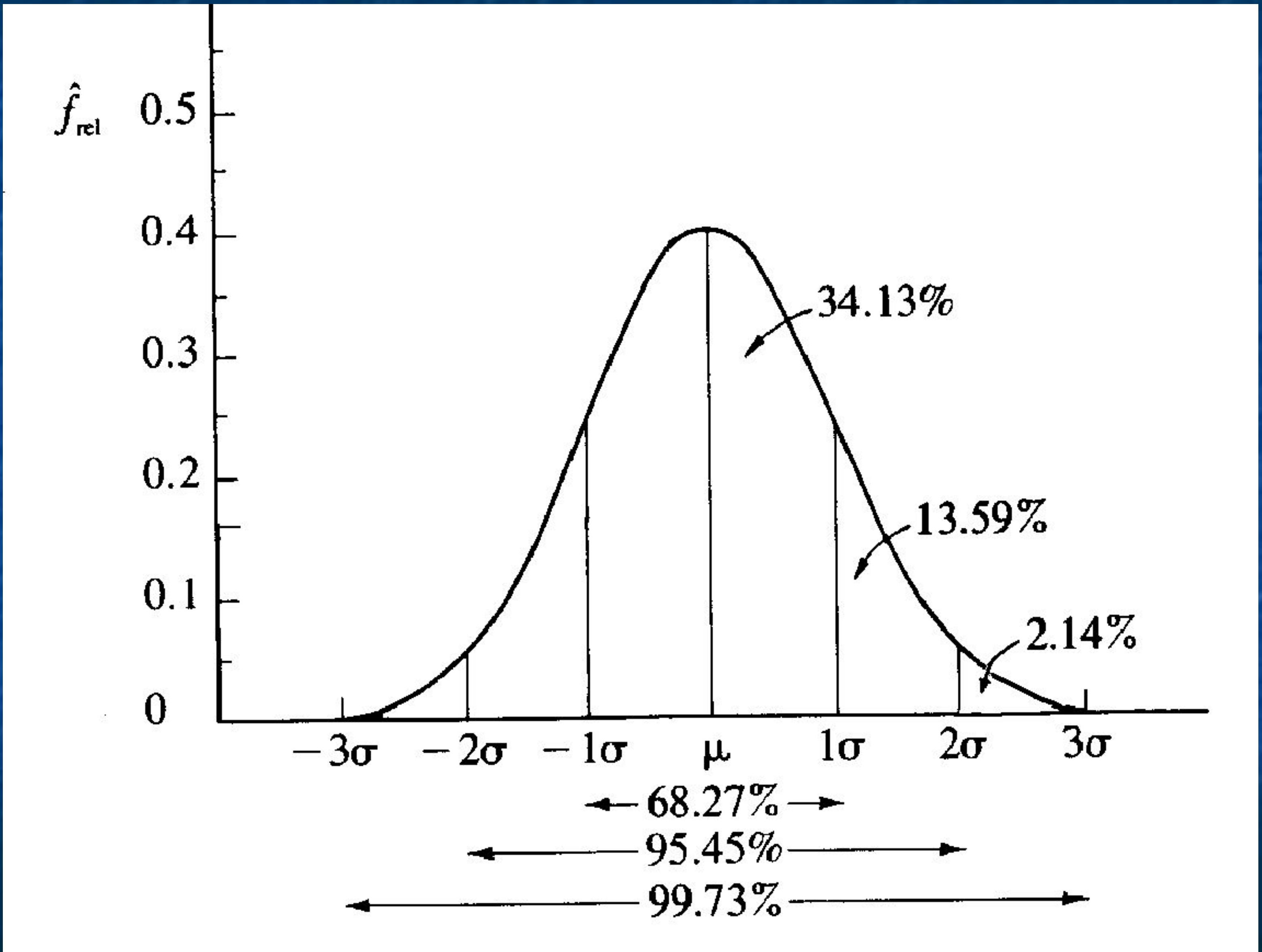
1) У нормального распределения арифметическая средняя, мода и медиана совпадают.



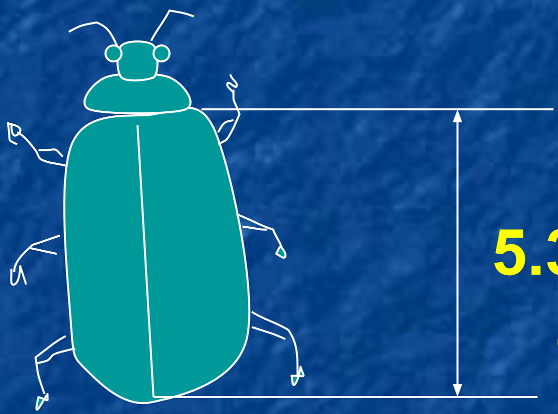
2) Правило трех сигм:

Подавляющее большинство значений нормально распределенного признака (99,73%) укладывается в интервал ±3 стандартных отклонения относительно среднего значения (т.е. $\mu \pm 3\sigma$)

Иллюстрация правила трех сигм:

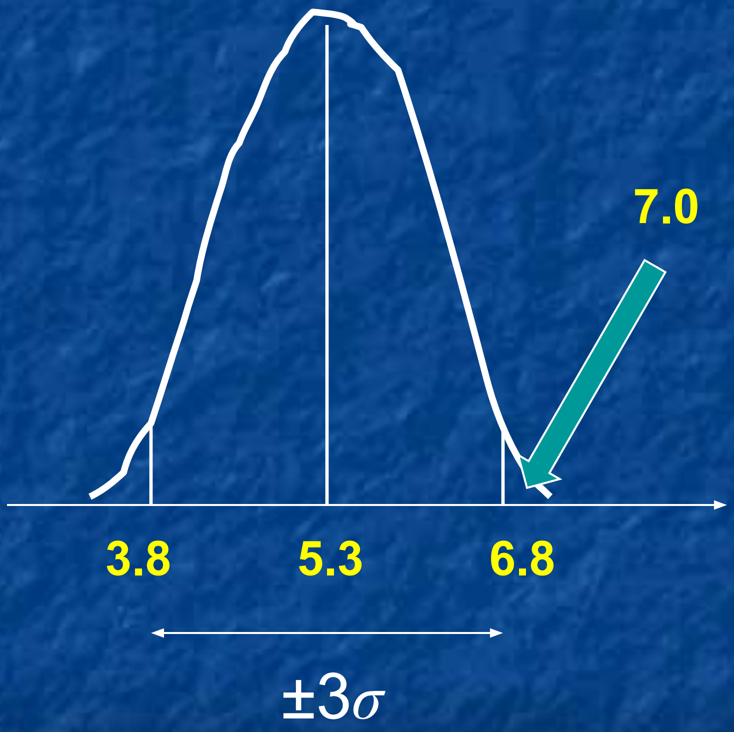


Пример использования правила трех сигм



5.3 ± 0.5 мм

7.0 мм
???



4.4. Биномиальное распределение

«Эксперимент со студентами»:

<u>День 1</u>: «отлавливаются» по 1 студенту;

<u>День 2</u>: группы по 2 студента;

<u>День 3</u>: по 3 студента.

p **– вероятность «поимки» белорусского студента (Б);**

q **- вероятность «поимки» иностранного студента (И).**

«Эксперимент со студентами»,

День 1:

Вероятность встретить хоть какого-нибудь студента равна:

$$p+q = 1$$

«Эксперимент со студентами»,

День 2:

$$p^2 + 2pq + q^2$$

P(ББ)

P(БИ) + P(ИБ)

P(ИИ)

«Эксперимент со студентами»,

День 3:

$$p^3 + 3p^2q + 3pq^2 + q^3, \text{ где}$$

- p^3 = Р(БББ);
- $3p^2q$ = Р(ББИ) + Р(БИБ) + Р(ИББ)
- $3pq^2$ = Р(БИИ) + Р(ИБИ) + Р(ИИБ)
- q^3 = Р(ИИИ)

«Эксперимент со студентами», обобщенно:

$(p + q)^n$ – бином Ньютона

- p – вероятность ожидаемого события;
- q – вероятность противоположного события;
- n – число испытаний (=объем выборки).

Пример:

P(Б) = 0,75

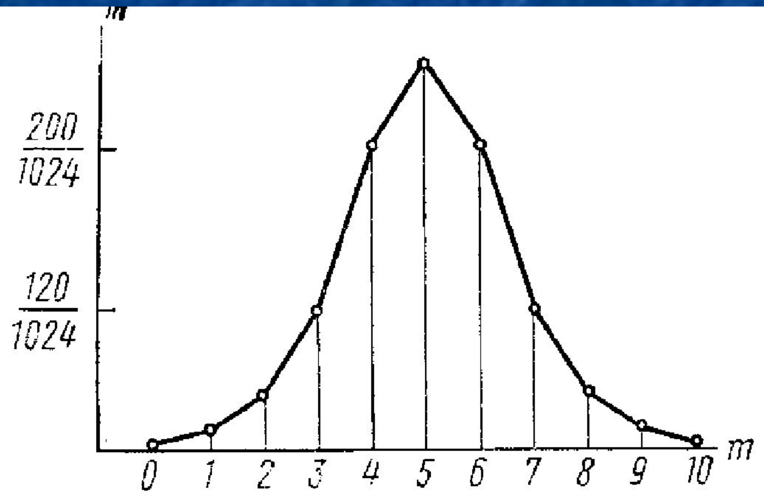
P(И) = 0,25

⇒ **вероятность того, что в группе из трех человек все окажутся беларусами: $(0,75)^3 = 0,422$**

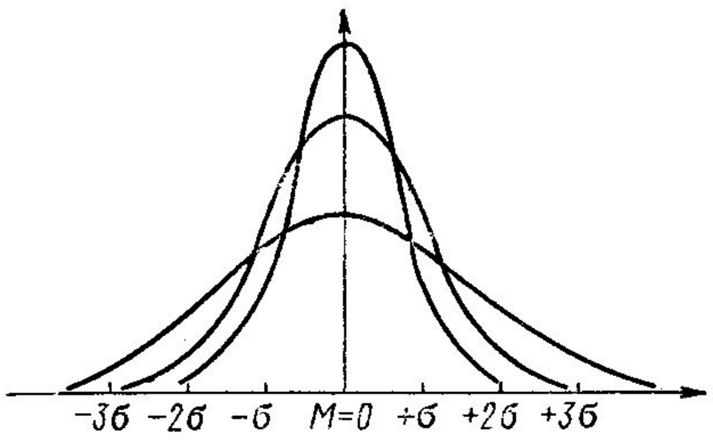
⇒ **аналогично, вероятность встречи трех иностранцев: $(0,25)^3 = 0,016$**

(!) Биномиальный закон описывает изменчивость только альтернативных признаков (*белорус/иностранец, черное/белое и т.п.*)

Биномиальное распределение определяется двумя параметрами: *p* и *n*



Кривая распределения вероятностей двучлена $(0,5+0,5)^{10}$



p=q

Расчет статистических параметров при биномиальном распределении:

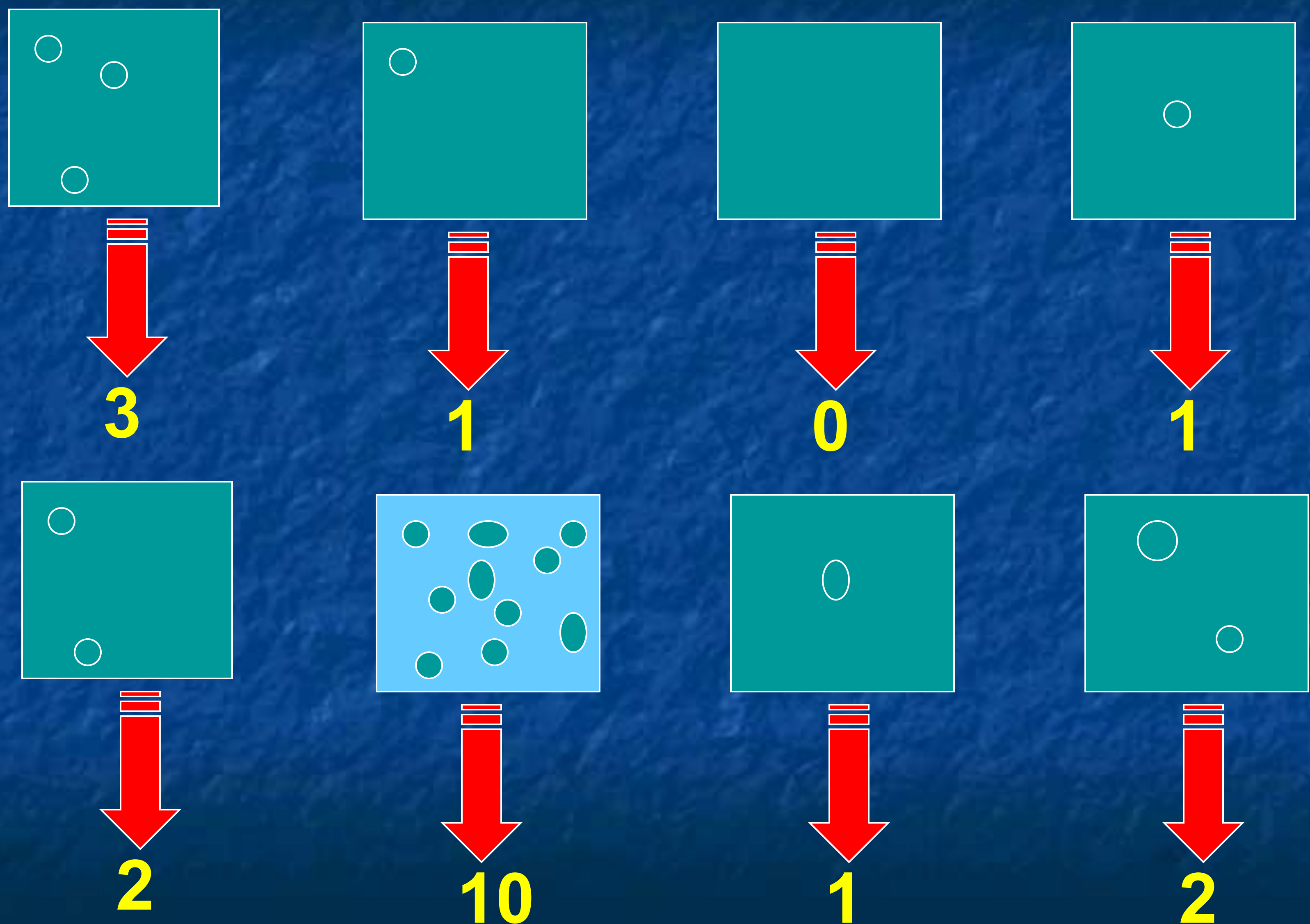
Средняя = np

Дисперсия = npq

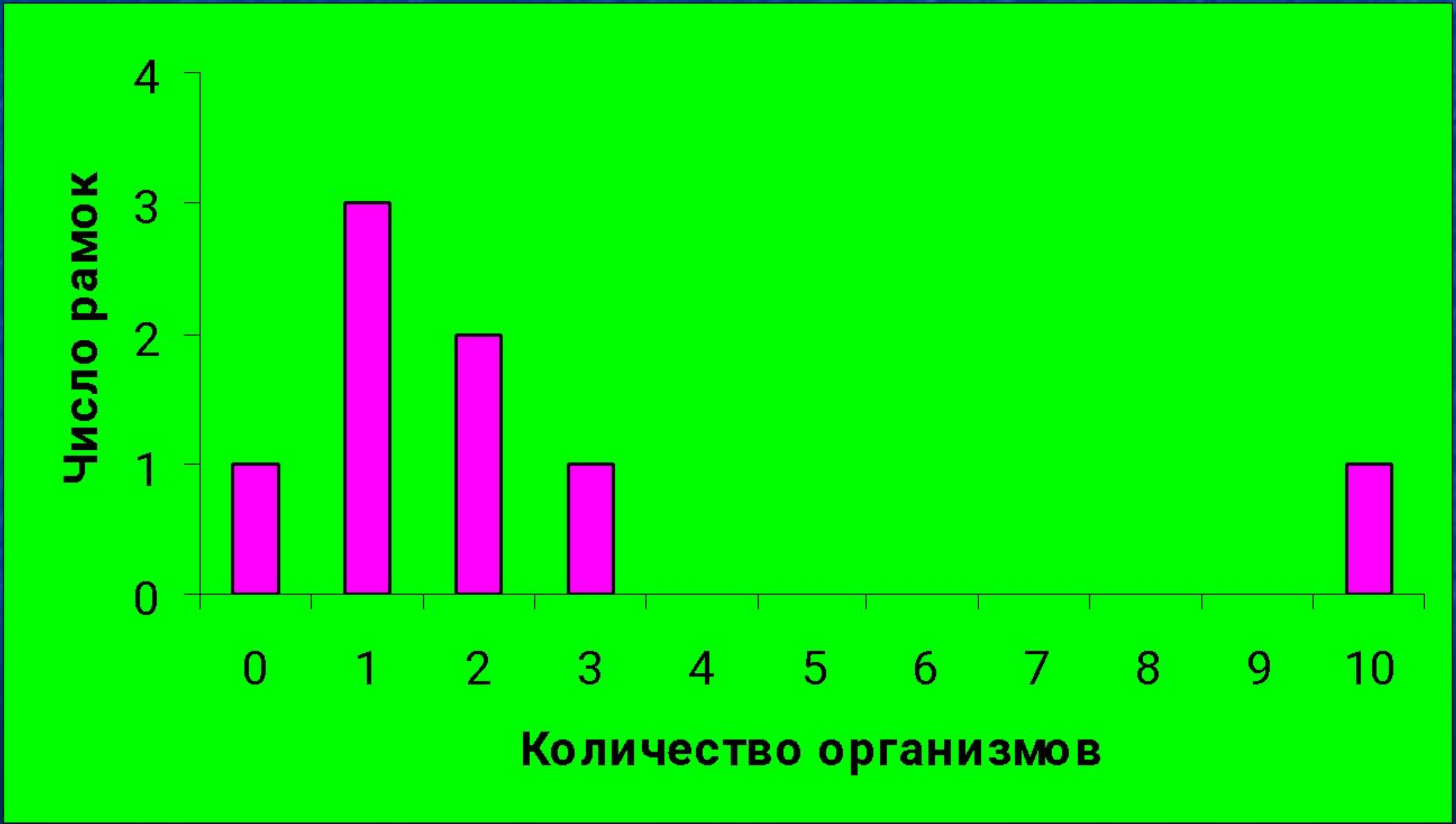
Стандартное отклонение = $\sqrt{npq}$

4.5. Негативное биномиальное распределение

Метод учетных площадок



Негативное биномиальное распределение



Математическое описание негативного биномиального распределения

$$(p + q)^{-k}$$

p **–** вероятность обнаружения определенного числа организмов в рамке

k – параметр, характеризующий степень агрегированности организмов (чем меньше *k*, тем более агрегированны организмы)

4.6. Закон Пуассона

Распределение Пуассона – предельный случай биномиального распределения – проявляется, когда:

- n велико (например, $(p + q)^{100}$);
- Вероятность ожидаемого события p мала (например, 0.01).

Примеры случайных событий:

- Возникновение летальных мутаций у бактерий за одну генерацию;
- Заболевание гриппом летом;
- Рождение тройни;
- Встреча большого числа организмов в учетной рамке;
- и т.п…

Математическое выражение закона Пуассона:

$$P_n(m) = \frac{a^m}{m!e^a}$$

m – частота ожидаемого события в *n* независимых испытаниях;

a ≅ *np* – средняя частота наступления события;

e = 2,7183 – основание натурального логарифма.

Распределение Пуассона определяется только одним параметром: *а* (средней)

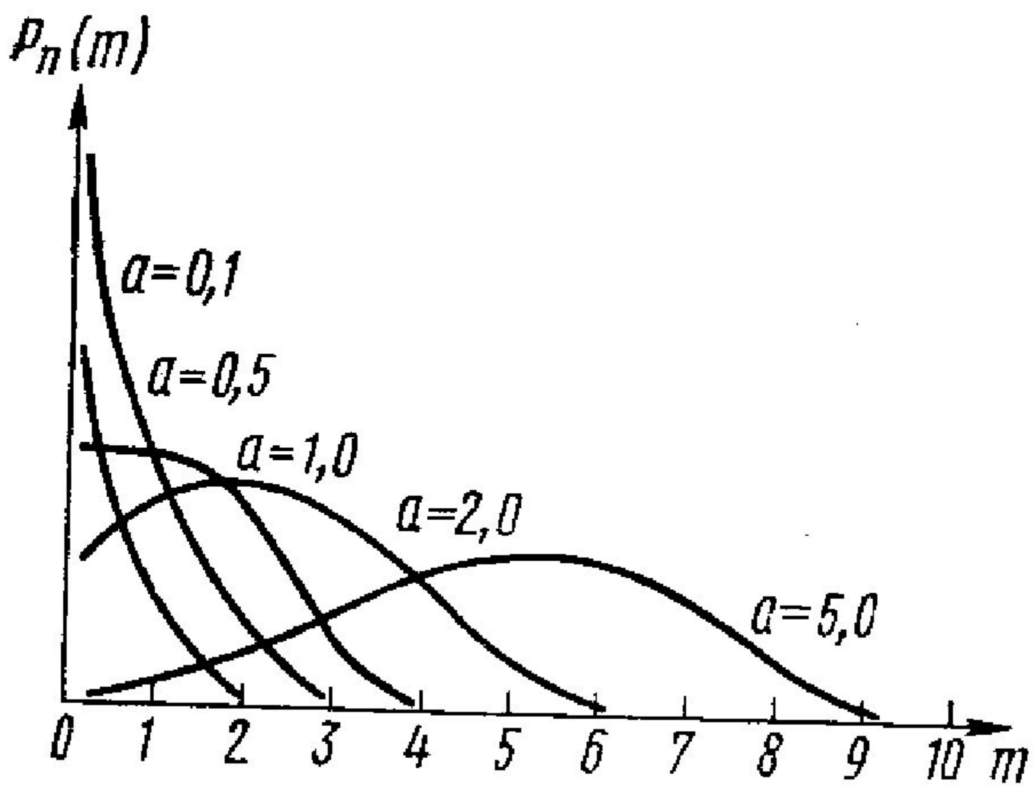


График функции $P_n(m) = \frac{a^m}{m!} \cdot e^{-a}$ для разных значений a

Статистические параметры при пуассоновском распределении:

Средняя = дисперсия = ***np***

Стандартное отклонение = ***√np***

Пьер Симон де Лаплас

(1749-1827)

«Замечательно, что наука, которая началась с рассмотрения азартных игр, обещает стать наиболее важным объектом человеческого знания...

Ведь большей частью важнейшие жизненные вопросы являются на самом деле лишь задачами из теории вероятностей»