Лекция 4. Виды распределений

Биномиальное, пуассоновское, геометрическое распределения, их производящие функции и числовые характеристики. Непрерывная случайная величина, функция распределения, плотность распределения, свойства плотности. Числовые характеристики непрерывных случайных величин. Основные непрерывные распределения (равномерное, показательное).

4.1. Биномиальное распределение

Пусть проведено n независимых испытаний с вероятностью p появления события A в каждом испытании (испытания Бернулли). Обозначим ξ – случайную величину, равную числу появлений события A в n испытаниях. По формуле Бернулли

$$P\{\xi = m\} = P(m) = C_n^m p^m q^{n-m}, \text{ где } q = 1 - p,\ m = 0{,}1,\ \ldots, n. \quad (4.1)$$

Определение 4.1. *Распределение дискретной случайной величины, задаваемое нижеприведенной таблицей, называется биномиальным.*

ξ	0	1	2	$\ldots$	k	$\ldots$	n
p	q^n	npq^{n-1}	$C_n^2 p^2 q^{n-2}$	$\ldots$	$C_n^k p^k q^{n-k}$	$\ldots$	p^k

Биномиальное распределение определяется двумя параметрами n и p.

Докажем, что сумма всех вероятностей равна 1.

Действительно, в соответствии с биномом Ньютона:

$$(p+q)^n = \sum_{m=0}^{n} C_n^m p^m q^{n-m}.$$

Но, с другой стороны: $(p+q)^n = \big(p + (1-p)\big)^n = 1$.

Найдем математическое ожидание и дисперсию такой случайной величины. Для этого представим ξ в виде суммы n независимых случайных величин ξ_i $(i = 1, 2,\ \ldots, n)$, таких, что $\xi_i = 1$, если в i-м испытании появилось событие A, и $\xi_i = 0$, если в i-м испытании появилось событие $\bar{A}$. Очевидно, что

$$\xi = \xi_1 + \ \ldots + \xi_n. \quad (4.2)$$

В этой сумме столько единиц, сколько раз появилось событие A в n испытаниях; остальные слагаемые равны нулю.

Распределение каждой из случайных величин ξ_i задаётся таблицей:

ξ_i	1	0
p	p	$1-p$

Очевидно, что

$$M(\xi_i) = 1 \cdot p + 0 \cdot (1-p) = p,$$

$$D(\xi_i) = M(\xi_i^2) - \big(M(\xi_i)\big)^2 = 1^2 \cdot p + 0^2 \cdot (1-p) - p^2 = p - p^2 =$$

$= p(1-p) = p \cdot q.$

Пользуясь свойствами математического ожидания и дисперсии, получаем два независимых слагаемых в формуле (4.2):

$$M(\xi) = M(\xi_1) + \ldots + M(\xi_n) = n \cdot p,$$

$$D(\xi) = D(\xi_1) + \ldots + D(\xi_n) = n \cdot p \cdot q.$$

Итак, для биномиально распределённой случайной величины ξ получим:

$$M(\xi) = np; \quad D(\xi) = npq. \tag{4.3}$$

Пример 4.1. *Монета брошена 4 раза. Написать закон распределения, найти математическое ожидание и дисперсию числа выпадений орла.*

►Найдём вероятности выпадения орла по формуле Бернулли при $n = 4$, $p = 0{,}5$:

$$P_4(0) = 0{,}5^4 \approx 0{,}0625; \quad P_4(1) = 4 \cdot 0{,}5 \cdot 0{,}5^3 = 0{,}25;$$

$$P_4(2) = C_4^2 \cdot 0{,}5^2 \cdot 0{,}5^2 \approx 0{,}375;$$

$$P_4(3) = p_4(1) = 0{,}25; \quad P_4(4) = p_4(0) \approx 0{,}0625.$$

Искомый закон распределения задаётся таблицей:

ξ	0	1	2	3	4
p	0,0625	0,25	0,375	0,25	0,0625

По формулам (4.3) находим:

$$M(\xi) = n \cdot p = 4 \cdot 0{,}5 = 2; \quad D(\xi) = nqp = 4 \cdot 0{,}5 \cdot 0{,}5 = 1. \blacktriangleright$$

Ответ: $M(\xi) = 2,\ D(\xi) = 1.$

Пример 4.2. *Производится двадцать выстрелов по мишени. Вероятность попадания при первом выстреле равно* 0,1, *а при каждом последующем выстреле производится корректировка прицела, поэтому вероятность попадания увеличивается на 10%. Написать закон распределения, найти математическое ожидание и дисперсию числа попаданий в мишень.*

►Подсчитываем вероятности попаданий при каждом выстреле. Для этого используем формулу сложных процентов: $p_k = 0{,}1(1 + 0{,}1)^k$, $k = \overline{0, 20}$. Получаем следующие значения массивов попаданий в цель p и промахов $q = 1 - p$:
$(p)[0.1, 0.11, 0.121, 0.1331, \cdots]$

Применяем формулу (2.14) для $n = 20$ и полученных массивов p и q.

Для решения задачи используем Maxima-программу. На рис. 18 представлен график функции распределения данной задачи.

```
kill(all)$  fpprintprec:4$N:20$
p:makelist(0.1*(1+0.1)^k,k,0,N-1);q:1-p;
P:product((q[k]+p[k]*z),k,1,N);
Fi:expand(P);
K:makelist(coeff(Fi,z^n),n,0,N)$  K[1]:coeff(Fi,z,0)$K;
s:sum(K[i],i,1,N+1);
plot2d([discrete, K], [x,1,14],[style,points])$
(p) [0.1,0.11,0.121,0.1331,0.1464,0.1611,0.1772,0.1949,
     0.2144,0.2358,0.2594,0.2853,0.3138,0.3452,0.3797,
     0.4177,0.4595,0.5054,0.556,0.6116]
(q) [0.9,0.89,0.879,0.8669,0.8536,0.8389,0.8228,0.8051,
     0.7856,0.7642,0.7406,0.7147,0.6862,0.6548,0.6203,
     0.5823,0.5405,0.4946,0.444,0.3884]
(P) (0.1*z+0.9)*(0.11*z+0.89)*(0.121*z+0.879)*(0.1331*z+0.8669)*
(0.1464*z+0.8536)*(0.1611*z+0.8389)*(0.1772*z+0.8228)*
*(0.1949*z+0.8051)*(0.2144*z+0.7856)*(0.2358*z+0.7642)*
*(0.2594*z+0.7406)*(0.2853*z+0.7147)*(0.3138*z+0.6862)*
*(0.3452*z+0.6548)*(0.3797*z+0.6203)*(0.4177*z+0.5823)*
(0.4595*z+0.5405)*(0.5054*z+0.4946)*(0.556*z+0.444)*
*(0.6116*z+0.3884)
(Fi) 7.322*10^-13*z^20+5.392*10^-11*z^19+1.841*10^-9*
*z^18+3.87*10^-8*z^17+5.616*10^-7*z^16+5.976*10^-6*z^15+
```

```
*4.835*10^-5*z^14+3.044*10^-4*z^13+0.001514*z^12+0.005997*
*z^11+0.01903*z^10+0.04841*z^9+0.09848*z^8+0.1592*
z^7+0.2025*z^6+0.1993*z^5+0.1481*z^4+0.08008*z^3+
+0.02961*z^2+0.006671*z+6.883*10^-4
(%o9) [6.883*10^-4,0.006671,0.02961,0.08008,0.1481,0.1993,
0.2025,0.1592,0.09848,0.04841,0.01903,0.005997,0.001514,
3.044*10^-4,4.835*10^-5,5.976*10^-6,5.616*10^-7,3.87*10^-8,
1.841*10^-9,5.392*10^-11,7.322*10^-13]
(%o11) 1.0
```

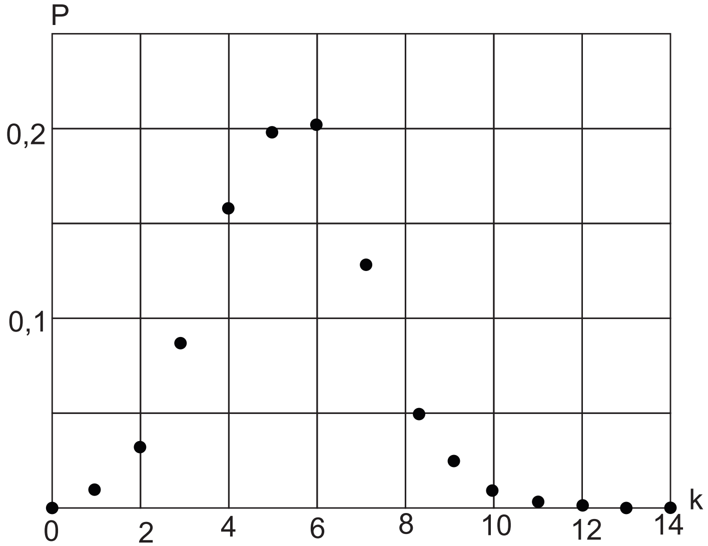

Рис. 18. *Распределение для примера 4.2*

4.2. Распределение Пуассона

Пусть в испытаниях Бернулли $n \to \infty$, $p \to 0$, так, что $np \to \lambda$. Тогда, как отмечалось ранее, вероятность $P_n(m)$ приближённо определяется с помощью формулы Пуассона:

$$P\{\xi = m\} = \frac{\lambda^m}{m!}e^{-\lambda}, \quad \lambda > 0. \tag{4.4}$$

Определение 4.2. *Распределение дискретной случайной величины, задаваемое формулой (4.4), называется распределением Пуассона или пуассоновским распределением.*

Запишем закон распределения Пуассона в виде таблицы:

ξ	0	1	2	...	m	...
p	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!}e^{-\lambda}$	...	$\frac{\lambda^m}{m!}e^{-\lambda}$	...

Распределение Пуассона определяется одним параметром λ.

Докажем, что сумма всех вероятностей равна 1.

Действительно, используя разложение в ряд Тейлора для e^{λ}, получим:

$$\sum_{m=0}^{\infty} \frac{\lambda^m}{m!}e^{-\lambda} = e^{-\lambda}\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

Найдём математическое ожидание и дисперсию такой случайной величины.

$$M(\xi) = \sum_{m=0}^{\infty} m\frac{\lambda^m}{m!}e^{-\lambda} = \sum_{m=1}^{\infty} m\frac{\lambda^m}{m!}e^{-\lambda} = \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!}e^{-\lambda} =$$

$$= \lambda \cdot e^{-\lambda}\sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = \lambda \cdot e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda.$$

Примем без доказательства, что $D(\xi) = \lambda$.

Итак, для случайной величины, имеющей распределение Пуассона, получим:

$$M(\xi) = D(\xi) = \lambda. \tag{4.5}$$

4.3. Геометрическое распределения

Пусть производится ряд независимых испытаний («попыток») для достижения некоторого результата (события A), и при каждой попытке событие A может появиться с вероятностью p. Тогда число попыток ξ до появления события A, включая удавшуюся, является дискретной случайной величиной, возможные значения которой принимают значения: $m = 1, 2, \ldots, m, \ldots$ Вероятности их по теореме умножения вероятностей для независимых событий равны

$$P(\xi = m) = pq^{m-1}, \qquad m = 1, 2, \ldots. \tag{4.6}$$

где $0 < p < 1, \quad q = 1 - p$.

Ряд распределения ξ имеет вид

ξ	1	2	3	$\ldots$	m	$\ldots$
P	p	pq	pq^2	$\ldots$	pq^{m-1}	$\ldots$

Как видно, вероятности $P_m = P(\xi = m) = pq^{m-1}, \quad m = 1, 2, \ldots,$ образуют для ряда последовательных значений бесконечно убывающую геометрическую прогрессию с первым членом p и знаменателем q (потому распределение и называется геометрическим). Сумма вероятностей возможных значений случайной величины будет равна

$$S = p + pq + pq^2 + \ldots + p + pq^{m-1} + \ldots = \frac{p}{1-q} = 1.$$

Примеры случайных величин, распределенных по геометрическому закону: число выстрелов до первого попадания, число испытаний устройства до первого отказа, число бросаний монеты до первого выпадения герба (или решки) и т.п. Найдем математическое ожидание и дисперсию при геометрическом распределении:

$$M(\xi) = 1 \cdot p + 2pq + \ldots + kq^{k-1} + \ldots = \sum_{k=1}^{\infty} kp^{k-1} = p \sum_{k=1}^{\infty} \frac{d(q^k)}{dq} =$$

$$= p\frac{d}{dq} \sum_{k=1}^{\infty} q^k = p\frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{p}{(1-q)^2} = \frac{1}{p},$$

так как $\sum\limits_{k=1}^{\infty} q^k = \dfrac{1}{1-q}$ — сумма членов бесконечно убывающей геометрической прогрессии с первым членом $b_1 = 1$ и знаменателем $q < 1$.

Следовательно,

$$D(\xi) = M(\xi^2) - m^2(\xi) = \frac{2q}{p^2} + \frac{1}{p} - \left(\frac{1}{p}\right)^2 = \frac{2q - 1 + p}{p^2} = \frac{q}{p^2}.$$

4.4. Гипергеометрическое распределения

С гипергеометрическими распределениями мы встречались когда решали задачу о выборке. Гипергеометрическое распределение широко используется в практике статистического приемочного контроля качества продукции, в задачах организации выборочных обследований и др.

Таблица распределения имеет вид:

ξ	0	1	2	$\ldots$	l
p	$\frac{C_L^0 C_{K-L}^k}{C_K^k}$	$\frac{C_L^1 C_{K-L}^{k-1}}{C_K^k}$	$\frac{C_L^2 C_{K-L}^{k-2}}{C_K^k}$	$\ldots$	$\frac{C_L^l C_{K-L}^0}{C_K^k}$

Здесь $k \leqslant K$, $l = min(k; L)$, $L \leqslant K$ и сумма всех вероятностей равна единице.

Типичное толкование: случайная величина ξ равна числу белых шаров, попавших в выборку без возвращения k шаров из урны, содержащей K шаров, из которых L белых.

$$P(\xi = m) = \frac{C_L^m \cdot C_{K-L}^k}{C_K^k}, \quad k = 0, 1, 2, \ldots, l.$$

$$M(\xi) = k \cdot \frac{L}{K}.$$

Рассмотренные распределения являются распределениями дискретных случайных величин. Далее рассмотрим некоторые распределения непрерывных случайных величин.

4.5. Равномерное распределение

Определение 4.3. *Распределение непрерывной случайной величины называется равномерным на $[a; b]$, если плотность распределения* *постоянна и отлична от 0 на этом отрезке и равна нулю вне его:*

$$f(x) = \begin{cases} C & \text{\textit{при}} \quad x \in [a;b], \\ 0 & \text{\textit{при}} \quad x \notin [a;b]. \end{cases}$$

Используя 6 свойство плотности распределения (п. 3.18), найдём константу C.

$$\int\limits_{-\infty}^{+\infty} f(x)dx = 1 \implies \int\limits_{a}^{b} Cdx = 1 \implies$$

$$\implies Cx\Big|_a^b = 1 \implies C(b-a) = 1 \implies C = \frac{1}{b-a}.$$

Итак, плотность равномерно распределённой на $[a;b]$ случайной величины определяется по формуле:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{при} \quad x \in [a;b], \\ 0 & \text{при} \quad x \notin [a;b]. \end{cases} \tag{4.7}$$

С помощью свойства 4 плотности (п. 3.18) найдем функцию распределения:

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt.$$

$$\begin{array}{lll} \text{При} & x < a & F(x) = \displaystyle\int\limits_{-\infty}^{x} 0dt = 0; \\ \text{при} & a \leqslant x \leqslant b & F(x) = \displaystyle\int\limits_{-\infty}^{a} 0dt + \int\limits_{a}^{x} \frac{1}{b-a}dt = \frac{x-a}{b-a}; \\ \text{при} & x > b & F(x) = \displaystyle\int\limits_{-\infty}^{a} 0dt + \int\limits_{a}^{b} \frac{1}{b-a}dt + \int\limits_{b}^{x} 0dt = \frac{b-a}{b-a} = 1. \end{array}$$

Итак, мы получили функцию распределения равномерно распределённой на $[a;b]$ случайной величины:

$$F(x) = \begin{cases} 0 & \text{при} \quad x < a, \\ \dfrac{x-a}{b-a} & \text{при} \quad a \leqslant x \leqslant b, \\ 1 & \text{при} \quad b < x. \end{cases} \tag{4.8}$$

Равномерное распределение определяется двумя параметрами a и b. Графики плотности и функции распределения равномерной на [a;b] случайной величины представлены на рис. 19.

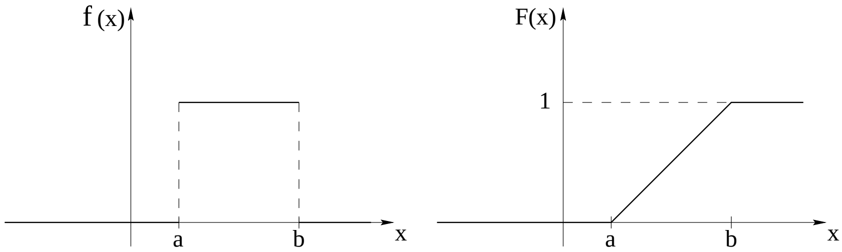

Рис. 19. *Плотность и функция распределения равномерного распределения*

Найдём математическое ожидание и дисперсию:

$$M(\xi) = \int\limits_{-\infty}^{+\infty} xf(x)dx = \int\limits_{a}^{b} \frac{x}{b-a}dx = \frac{1}{b-a} \cdot \frac{x^2}{2}\bigg|_a^b = \frac{b^2-a^2}{2\cdot(b-a)} = \frac{a+b}{2}$$

$$D(\xi) = \int\limits_{-\infty}^{+\infty} x^2 f(x)dx - \frac{(a+b)^2}{4} = \int\limits_{a}^{b} \frac{x^2}{b-a}dx - \frac{(a+b)^2}{4} =$$

$$= \frac{1}{b-a} \cdot \frac{x^3}{3}\bigg|_a^b - \frac{(a+b)^2}{4} = \frac{(b-a)^3}{3\cdot(b-a)} - \frac{(a+b)^2}{4} =$$

$$= \frac{4\cdot(a^2+ab+b^2) - 3\cdot(a+b)^2}{12} = \frac{(b-a)^2}{12}.$$

Итак, для равномерно распределённой на $[a;b]$ случайной величины получим:

$$M(\xi) = \frac{a+b}{2}; \quad D(\xi) = \frac{(b-a)^2}{12}. \tag{4.9}$$

Замечание 4.1. *Найдём $P\{x \leqslant \xi < x+\Delta x\}$ при условии, что $a \leqslant x < x+\Delta x \leqslant b$. Пользуясь свойством 5 плотности (п. 3.18), получаем:*

$$P\{x \leqslant \xi < x+\Delta x\} = \int\limits_{x}^{x+\Delta x} f(x)dt = \int\limits_{x}^{x+\Delta x} \frac{1}{b-a}dt = \frac{x+\Delta x - x}{b-a} = \frac{\Delta x}{b-a}.$$

Как видим, эта вероятность не зависит от x, т.е. от положения промежутка внутри $[a;b]$, а только от длины промежутка Δx. Этим объясняется название распределения — равномерное. Вероятность распределена «равномерно» по отрезку $[a;b]$ (плотность постоянна). Очевидно, что в этом случае среднее значение случайной величины равно середине отрезка: $M(\xi) = \frac{a+b}{2}$.

Пример 4.3. *Плотность распределения постоянна на отрезке $[0;4]$ и равна нулю вне его. Найти плотность и функцию распределения, математическое ожидание и дисперсию.*

►В соответствии с определением 4.3 эта случайная величина имеет равномерное распределение на отрезке $[0;4]$. Следовательно:

$$f(x) = \begin{cases} \dfrac{1}{4} & \text{при} \quad x \in [0;4], \\ 0 & \text{при} \quad x \notin [0;4], \end{cases} \qquad F(x) = \begin{cases} 0 & \text{при} \quad x < 0, \\ \dfrac{x}{4} & \text{при} \quad 0 \leqslant x \leqslant 4, \\ 1 & \text{при} \quad x > 4, \end{cases}$$

$$M(\xi) = 2; \quad D(\xi) = \frac{(4-0)^2}{12} = \frac{4}{3} \approx 1{,}333.$$

Ответ $M(\xi) = 2;\ D(\xi) = \dfrac{4}{3} \approx 1{,}333.$

4.6. Экспоненциальное распределение

Определение 4.4. *Распределение непрерывной случайной величины называется экспоненциальным (показательным), если плотность распределения имеет вид:*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \textit{при} \quad x \geqslant 0, \\ 0 & \textit{при} \quad x < 0, \quad \textit{где } \lambda > 0. \end{cases} \tag{4.10}$$

Экспоненциальное распределение определяется одним параметром $\lambda > 0$.

Найдем функцию распределения:

$$\text{при } x \geqslant 0 \quad F(x) = \int\limits_{-\infty}^{x} f(t)dt = \int\limits_{0}^{x} \lambda e^{-\lambda t}dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x};$$

$$\text{при } x < 0 \quad F(x) = \int\limits_{-\infty}^{x} f(t)dt = \int\limits_{-\infty}^{x} 0dt = 0.$$

Итак, функция распределения имеет вид:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{при} \quad x \geqslant 0, \\ 0 & \text{при} \quad x < 0. \end{cases} \tag{4.11}$$

Графики плотности и функции распределения экспоненциального распределения представлены на рис. 20.

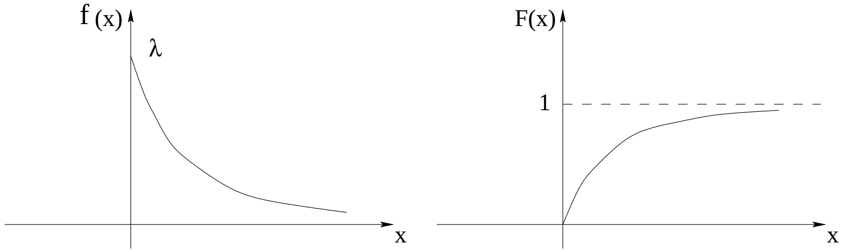

Рис. 20. *Плотность и функция распределения экспоненциального распределения*

Найдём математическое ожидание и дисперсию.

$$M(\xi) = \int\limits_{-\infty}^{+\infty} x f(x) dx = \int\limits_{0}^{+\infty} x\lambda e^{-\lambda x} dt = \lambda \int\limits_{0}^{+\infty} x e^{-\lambda x} dx =$$

$$= \left[\begin{array}{cc} u = x & du = dx \\ dv = e^{-\lambda x} & v = -\dfrac{1}{\lambda} e^{-\lambda x} \end{array} \right] = -x e^{-\lambda x} \Big|_0^{\infty} + \int\limits_0^{+\infty} e^{-\lambda x} dx =$$

$$= -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}.$$

Самостоятельно проведите выкладки и докажите, что:

$$D(\xi) = \int\limits_{-\infty}^{+\infty} x^2 f(x) dx - \frac{1}{\lambda^2} = \int\limits_0^{+\infty} x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Итак, для экспоненциально распределённой случайной величины получим:

$$M(\xi) = \frac{1}{\lambda}; \quad D(\xi) = \frac{1}{\lambda^2}. \tag{4.12}$$

Замечание 4.2. *Можно доказать, что если через независимые случайные промежутки времени* $\xi_1, \xi_2, \xi_3, \ldots$, *имеющие экспоненциальное распределение с параметром* λ, *происходит какое-либо событие (например, поступает вызов на телефонную станцию или приходит покупатель в магазин), то количество этих событий, произошедших за любой промежуток времени* t, *является случайной величиной, имеющей пуассоновское распределение с параметром* $a = \lambda t$.