Формулы по теории вероятностей

I. Случайные события

1. Основные формулы комбинаторики

а) перестановки $P_n = n! = 1 \cdot 2 \cdot 3 ... (n-1) \cdot n$.

б) размещения $A_n^m = n \cdot (n-1)...(n-m+1)$.

в) сочетания $C_n^k = \dfrac{A_n^k}{P_k} = \dfrac{n!}{(n-k)! \cdot k!}$.

2. Классическое определение вероятности.

$P(A) = \dfrac{m}{n}$, где m - число благоприятствующих событию A исходов, n - число всех элементарных равновозможных исходов.

3. Вероятность суммы событий

Теорема сложения вероятностей несовместных событий:
$P(A+B) = P(A) + P(B)$

Теорема сложения вероятностей совместных событий:
$P(A+B) = P(A) + P(B) - P(AB)$

4. Вероятность произведения событий

Теорема умножения вероятностей независимых событий:
$P(A \cdot B) = P(A) \cdot P(B)$

Теорема умножения вероятностей зависимых событий:
$P(A \cdot B) = P(A) \cdot P(B|A)$,
$P(A \cdot B) = P(B) \cdot P(A|B)$.

$P(A|B)$ - условная вероятность события A при условии, что произошло событие B,
$P(B|A)$ - условная вероятность события B при условии, что произошло событие A.

5. Формула полной вероятности

$P(A) = \sum_{k=1}^{n} P(H_k)P(A|H_k)$, где $H_1, H_2, ..., H_n$ - полная группа гипотез, то есть

$H_i \cdot H_j = \varnothing,\ i \neq j,\ \bigcup_{i=1}^{n} H_i = \Omega$, Ω - достоверное событие.

6. Формула Байеса (формула Бейеса). Вычисление апостериорных вероятностей гипотез

$$P(H_m|A) = \frac{P(H_m)P(A|H_m)}{\sum_{k=1}^{n} P(H_k)P(A|H_k)},\ m = 1,...,n,$$ где $H_1, H_2, ..., H_n$ - полная группа гипотез.

7. Формула Бернулли

$P_n(k)=C_n^k p^k(1-p)^{n-k}=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$ - вероятность появления события ровно k раз при n независимых испытаниях, p - вероятность появления события при одном испытании.

8. Наивероятнейшее число наступления события.

Наивероятнейшее число k_0 появления события при n независимых испытаниях:
$np-(1-p)\le k_0<np+p$, p - вероятность появления события при одном испытании.

9. Локальная формула Лапласа

$P_n(k)\approx\frac{1}{\sqrt{npq}}\varphi\left(\frac{k-np}{\sqrt{npq}}\right)$ - вероятность появления события ровно k раз при n независимых испытаниях, p - вероятность появления события при одном испытании, $q=1-p$.

10. Интегральная формула Лапласа

$P_n(m_1,m_2)\approx\Phi\left(\frac{m_2-np}{\sqrt{npq}}\right)-\Phi\left(\frac{m_1-np}{\sqrt{npq}}\right)$ - вероятность появления события не менее k_1 и не более k_2 раз при n независимых испытаниях, p - вероятность появления события при одном испытании, $q=1-p$.

11. Оценка отклонения относительной частоты от постоянной вероятности p:

$$P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\approx2\Phi\left(\varepsilon\sqrt{\frac{n}{p(1-p)}}\right).$$

II. Случайные величины

12. Ряд распределения дискретной случайной величины

x_i	x_1	x_2	…….	x_n
p_i	p_1	p_2	…….	p_n

Сумма вероятностей всегда равна 1. $\sum_{i=1}^{n}p_i=1$

13. Функция распределения (интегральная функция распределения)

Функция распределения случайной величины X определяется по формуле $F(x)=P(X<x)$. Это неубывающая функция, принимающая значения от 0 до 1. Если

задана плотность распределения $f(x)$, то функция распределения выражается как $F(x)=\int_{-\infty}^{x} f(t)\,dt$.

14. Плотность распределения (дифференциальная функция распределения)
Плотность распределения случайной величины X определяется по формуле $f(x)=F'(x)$. Существует только для непрерывной случайной величины. Для нее выполняется условие нормировки: $\int_{-\infty}^{\infty} f(x)\,dx=1$ (площадь под кривой равна 1).

15. Вероятность попадания случайной величины в заданный интервал
Может быть вычислена двумя способами:
1) через функцию распределения $P(\alpha<X<\beta)=F(\beta)-F(\alpha)$
2) через плотность распределения $P(\alpha<X<\beta)=\int_{\alpha}^{\beta} f(x)\,dx$

16. Математическое ожидание случайной величины
1) Для дискретной случайной величины X, заданной рядом распределения:
$$M(X)=\sum_{i=1}^{n} x_i\cdot p_i$$
1) Для непрерывной случайной величины X, заданной плотностью распределения :
$$M(X)=\int_{-\infty}^{\infty} f(x)\cdot x\,dx.$$

17. Дисперсия случайной величины
По определению дисперсия – это второй центральный момент: $D(X)=M\left(\left[X-M(X)\right]^2\right)=M\left(X^2\right)-\left(M(X)\right)^2$.
1) Для дискретной случайной величины X, заданной рядом распределения:
$$D(X)=\sum_{i=1}^{n}\left(x_i-M(X)\right)^2\cdot p_i=\sum_{i=1}^{n}x_i^2\cdot p_i-\left(M(X)\right)^2=\sum_{i=1}^{n}x_i^2\cdot p_i-\left(\sum_{i=1}^{n}x_i\cdot p_i\right)^2$$
1) Для непрерывной случайной величины X, заданной плотностью распределения :
$$D(X)=\int_{-\infty}^{\infty} f(x)\cdot\left(x-M(X)\right)^2 dx=\int_{-\infty}^{\infty} f(x)\cdot x^2\,dx-\left(M(X)\right)^2=\int_{-\infty}^{\infty} f(x)\cdot x^2\,dx-\left(\int_{-\infty}^{\infty} f(x)\cdot x\,dx\right)^2.$$

18. Среднее квадратическое отклонение случайной величины
$$\sigma=\sqrt{D(X)}$$

19. Начальный момент r–го порядка случайной величины

$\nu_r = M\left(X^r\right)$.

В частности, первый начальный момент – это математическое ожидание:

$\nu_1 = M\left(X^1\right) = M\left(X\right)$

20. Центральный момент r – го порядка случайной величины

$\mu_r = M\left(\left[X - M\left(X\right)\right]^r\right)$

В частности, второй центральный момент – это дисперсия:

$\mu_2 = M\left(\left[X - M\left(X\right)\right]^2\right) = D\left(X\right)$.

21. Асимметрия

$A_s = \dfrac{\mu_3}{\sigma^3}$

Коэффициент асимметрии положителен, если правый хвост распределения длиннее левого (правая часть кривой более пологая), и отрицателен в противном случае. Если распределение симметрично относительно математического ожидания, то его коэффициент асимметрии равен нулю.

22. Эксцесс

$E = \dfrac{\mu_4}{\sigma^4} - 3$

Коэффициент эксцесса нормального распределения равен нулю. Он положителен, если пик распределения около математического ожидания острый, и отрицателен, если пик гладкий.

III. Распределения случайных величин

21. Биномиальное распределение (дискретное)

X - количество «успехов» в последовательности из n независимых случайных экспериментов, таких что вероятность «успеха» в каждом из них равна p. $q = 1 - p$.

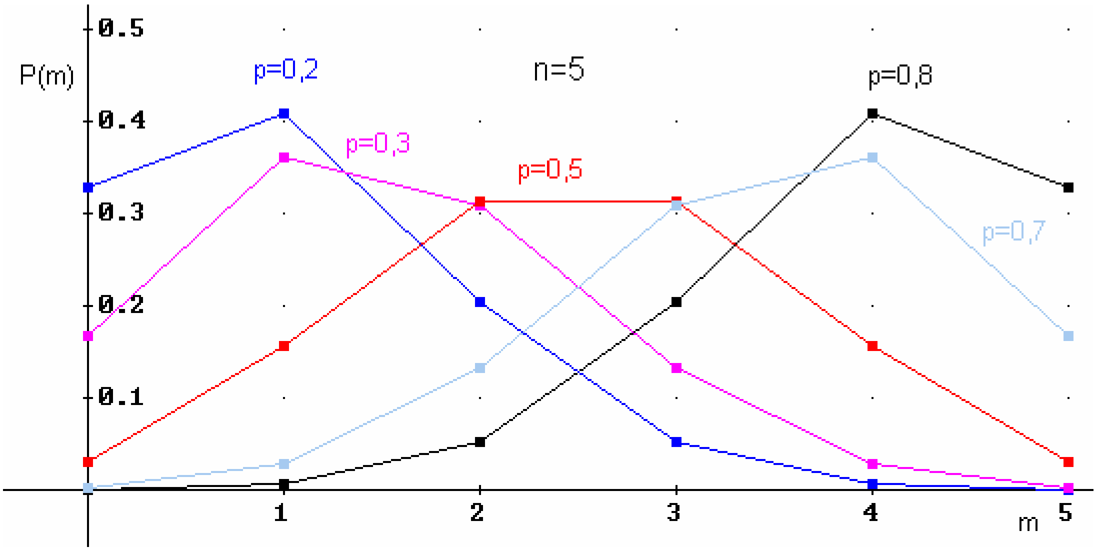
Закон распределения X имеет вид:

x_k	0	1	…..	k	…..	n
p_k	q^n	$n \cdot p \cdot q^{n-1}$		$C_n^k \cdot p^k \cdot q^{n-k}$		p^n

Здесь вероятности находятся по формуле Бернулли: $P\left(X = k\right) = C_n^k p^k \left(1 - p\right)^{n-k} = C_n^k p^k q^{n-k}$.

Характеристики: $M\left(X\right) = np$, $D\left(X\right) = npq$, $\sigma = \sqrt{npq}$

Примеры многоугольников распределения для $n = 5$ и различных вероятностей:



22. Пуассоновское распределение (дискретное)

Распределение Пуассона моделирует случайную величину, представляющую собой число событий, произошедших за фиксированное время, при условии, что данные события происходят с некоторой фиксированной средней интенсивностью и независимо друг от друга.

При условии $p \to 0,\ n \to \infty,\ np \to \lambda = const$ закон распределения Пуассона является предельным случаем биномиального закона. Так как при этом вероятность p события A в каждом испытании мала, то закон распределения Пуассона называют часто законом редких явлений.

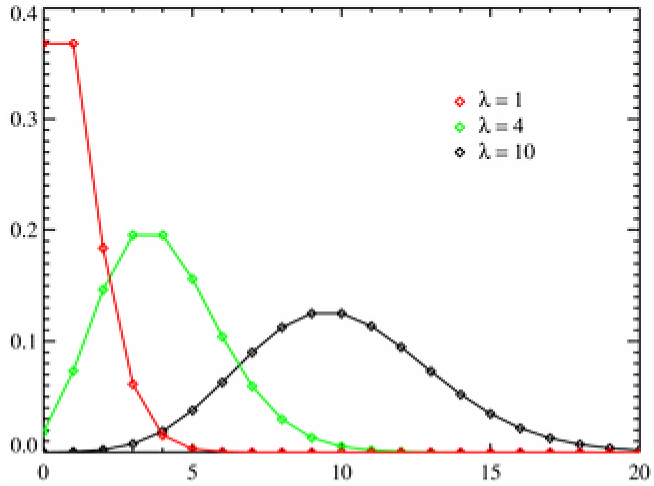
Ряд распределения:

x_k	0	1	…..	k	…..
p_k	$e^{-\lambda}$	$\lambda e^{-\lambda}$	…..	$\frac{\lambda^k}{k!}e^{-\lambda}$	…..

Вероятности вычисляются по формуле Пуассона: $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$.

Числовые характеристики: $M(X)=\lambda$, $D(X)=\lambda$, $\sigma=\sqrt{\lambda}$

Разные многоугольники распределения при $\lambda = 1; 4; 10$.



23. Показательное распределение (непрерывное)

Экспоненциальное или показательное распределение — абсолютно непрерывное распределение, моделирующее время между двумя последовательными свершениями одного и того же события.

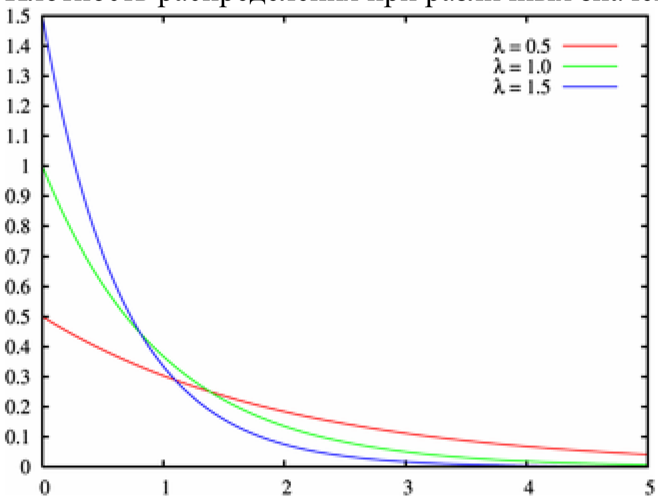
Плотность распределения:

$$p(x)=\begin{cases} 0, & x<0 \\ \lambda\cdot e^{-\lambda x}, & x\ge 0 \end{cases}$$

Где $\lambda>0$.

Числовые характеристики: $M(X)=\dfrac{1}{\lambda}$, $D(X)=\dfrac{1}{\lambda^2}$, $\sigma=\dfrac{1}{\lambda}$

Плотность распределения при различных значениях $\lambda>0$.



24. Равномерное распределение (непрерывное)

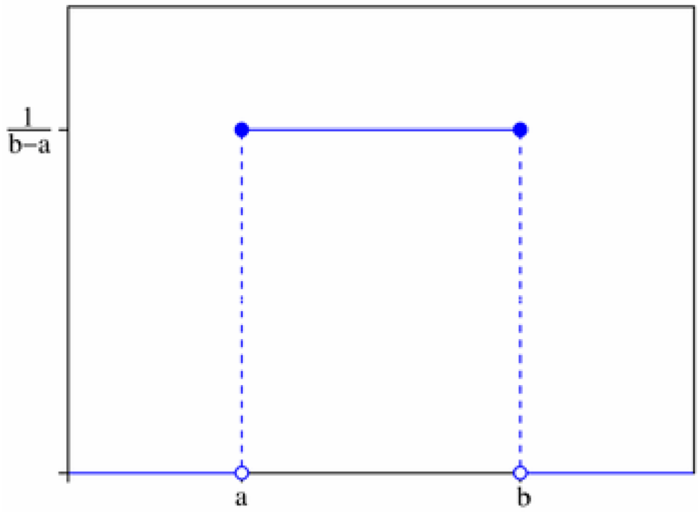
Равномерный закон распределения используется при анализе ошибок округления при проведении числовых расчётов (например, ошибка округления числа до целого распределена равномерно на отрезке [-0,5; 0,5]), в ряде задач массового обслуживания, при статистическом моделировании наблюдений, подчинённых заданному распределению.

Плотность распределения:

$$p(x)=\begin{cases} 0 & x\le a \\ \dfrac{1}{b-a} & a\le x\le b \\ 0 & x>b \end{cases}$$

Числовые характеристики: $M(X)=\dfrac{a+b}{2}$, $D(X)=\dfrac{(b-a)^2}{12}$, $\sigma=\dfrac{b-a}{2\sqrt{3}}$

График плотности вероятностей:



25. Нормальное распределение или распределение Гаусса (непрерывное)

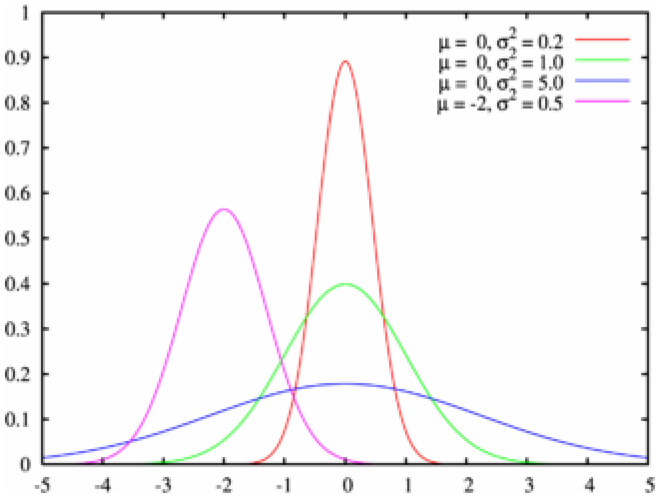
Нормальное распределение, также называемое распределением Гаусса, – распределение вероятностей, которое играет важнейшую роль во многих областях знаний, особенно в физике. Физическая величина подчиняется нормальному распределению, когда она подвержена влиянию огромного числа случайных помех. Ясно, что 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.

Плотность распределения:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$$

Числовые характеристики: $M(X)=a$, $D(X)=\sigma^2$, $\sigma=\sigma$

Пример плотности распределения:



Нормальный закон распределения случайной величины с параметрами $a=0$ и $\sigma=1$ называется стандартным или нормированным, а соответствующая нормальная кривая - стандартной или нормированной.

Функция Лапласа $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-t^2/2}dt$.

Вероятность попадания нормально распределенной случайной величины X в заданный интервал $(\alpha;\beta)$

$$P(\alpha < X < \beta)=\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right)$$

Вероятность отклонения нормально распределенной случайной величины X на величину δ от математического ожидания (по модулю).

$$P(|X-a|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right).$$

IV. Другие формулы

26. Неравенство Чебышева

$$P\left(|X-M(X)|<\varepsilon\right)\geq 1-\frac{D(X)}{\varepsilon^2}$$

27. Неравенство Маркова

$$P(X\leq\varepsilon)>1-\frac{MX}{\varepsilon}$$

28. Математическое ожидание функции одной случайной величины

$$M[\varphi(x)]=\sum_{i=1}^{n}\varphi(x_i)\cdot p_i$$

29. Корреляционный момент системы случайных величин X **и** Y

$$\mu_{XY} = M\left(\left[X - M\left(X\right)\right]\cdot\left[Y - M\left(Y\right)\right]\right) = M\left(X\cdot Y\right) - M\left(X\right)\cdot M\left(Y\right)$$

30. Коэффициент корреляции системы случайных величин X **и** Y

$$r_{XY} = \frac{\mu_{XY}}{\sigma_X \cdot \sigma_Y}$$

31. Пуассоновский поток событий

$$p_t(k) = \frac{(\lambda\, t)^k \cdot e^{-\lambda\, t}}{k!}$$