Здравствуйте!

Схема Бернулли.

Теорема об опытах. Биномиальное распределение

Схема Бернулли заключается в следующем. Проводится серия из n опытов. Предполагается, что выполнены условия

а) опыты независимы, то есть исход одного опыта не влияет на исход другого опыта;

б) в каждом опыте может наступить (или не наступить) некоторое случайное событие A, но вероятность $P(A) = p$ наступления этого события одна и та же в каждом опыте.

Требуется найти вероятность $P_n(m)$ того, что в серии из n опытов событие A наступит ровно m раз.

Теорема (об опытах). В схеме Бернулли вероятность $P_n(m)$ определяется равенством $P_n(m) = C_n^m p^m (1-p)^{n-m}$,

где число сочетаний C_n^m определяется равенством

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Для доказательства обозначим

A_i – событие, заключающееся в том, что в i-м опыте событие A наступает.

$B_n(m)$ – событие, заключающееся в том, что в серии из n опытов событие A наступит ровно m раз.

Тогда можно записать следующее разложение

$$B_n(m) = A_1 A_2 \ldots A_m \overline{A}_{m+1} \overline{A}_{m+2} \ldots \overline{A}_n + \overline{A}_1 A_2 \ldots A_m A_{m+1} \overline{A}_{m+2} \ldots \overline{A}_n + \ldots .$$

Отметим, что слагаемые в этой сумме попарно несовместные случайные события, поэтому

$$P\{B_n(m)\} = P\{A_1 A_2 \ldots A_m \overline{A}_{m+1} \overline{A}_{m+2} \ldots \overline{A}_n\} + P\{\overline{A}_1 A_2 \ldots A_m A_{m+1} \overline{A}_{m+2} \ldots \overline{A}_n\} + \ldots .$$

В силу независимости опытов, эту вероятность можно переписать в виде

$$P\{B_n(m)\} = p^m (1-p)^{n-m} + p^m (1-p)^{n-m} + \ldots .$$

Очевидно, что число слагаемых в этой сумме равно C_n^m, поэтому выполняется равенство

$$P_n(m) = P\{B_n(m)\} = C_n^m p^m (1-p)^{n-m}.$$

Теорема доказана.

Совокупность по всем $m = 0,1,2,\ldots,n$ вероятностей $P_n(m)$ называется **биномиальным распределением вероятностей.**

Для биномиального распределения вероятностей найдём то значение m, при котором вероятность $P_n(m)$ принимает наибольшее значение. Это значение $m = m^*$ называется **наивероятнейшим числом успехов**.

Рассмотрим отношение

$$\frac{P_n(m+1)}{P_n(m)} = \frac{n!\,p^{m+1}(1-p)^{n-m-1}}{(m+1)!(n-m-1)!}\frac{m!(n-m)!}{n!\,p^m(1-p)^{n-m}} = \frac{n-m}{m+1}\frac{p}{1-p}.$$

Найдём те значения m, при которых это отношение будет больше единицы.

Так как $\frac{n-m}{m+1}\frac{p}{1-p}>1$, следовательно $np-mp>m+1-mp-p$, поэтому $m<np+(p-1)$.

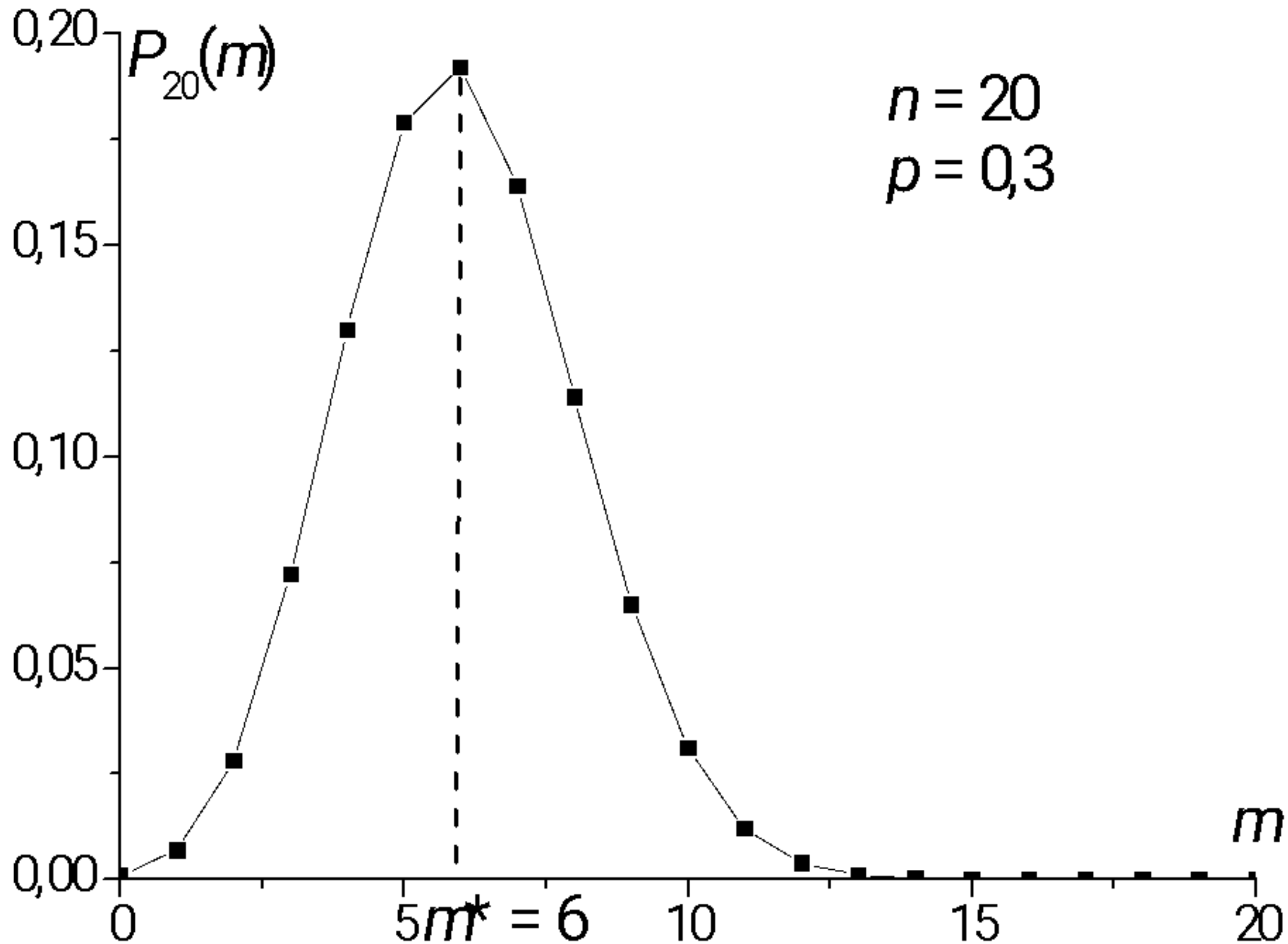
Итак, при $m<np+(p-1)$ выполняется неравенство $P_n(m+1)>P_n(m)$.

Очевидно, что при $m>np+(p-1)$ будет наоборот $P_n(m+1)<P_n(m)$.

Следовательно, максимальное значение вероятность $P_n(m)$ принимает при $m=m^*=]np+(p-1)[+1=]np+p[$,

где $]x[$ есть целая часть числа x, то есть наибольшее целое число, не превосходящее x.

Левее этого значения вероятности $P_n(m)$ монотонно возрастают с ростом m, а правее – убывают.



Зависимость $P_n(m)$ от m при $n = 20$ и $p = 0{,}3$

Задача 1. Монету бросают 6 раз. Найти вероятность того, что герб выпадет: а) менее двух раз; б) не менее двух раз.

Задача 2. Два равносильных противника играют в шахматы. Что вероятнее:

а) выиграть одну партию из двух или две партии из четырех?

б) выиграть не менее двух партий из четырех или не менее трех партий из пяти?

Задача 3. Укажите значения n, k, p, q. Напишите формулу вероятности $P_n(k)$.

1. Случайным образом называют 10 цифр. Какова вероятность того, что цифра 5 встретится ровно 7 раз?
2. «Хорошо», если наудачу выбранная карта из 36 – не бубновой масти. Карту каждый раз возвращают в колоду. Какова вероятность того, что ровно в 90 случаях из 200 таких вытаскиваний будет «плохо»?
3. Бросание кубика считается удачным, если выпадает 5 или 6 очков. Какова вероятность того, что 175 из 293 будут удачными?

Формулы Муавра-Лапласа

Несмотря на внешнюю простоту формулы для $P_n(m)$, применять её целесообразно лишь при небольших значениях n и m, в связи с необходимостью вычисления значений факториалов.

Поэтому для достаточно больших значений n, применяют другую формулу, которую даёт локальная теорема Муавра-Лапласа.

Локальная предельная теорема Муавра-Лапласа.

Теорема. В условиях схемы Бернулли обозначим

$$x = \frac{m - np}{\sqrt{np(1-p)}},$$

тогда в любом конечном интервале $a < x < b$, имеет место следующее равенство

$$\lim_{n \to \infty} \sqrt{np(1-p)} \cdot P_n(m) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}.$$

На практике её применяют в следующей форме

$$P_n(m) \approx \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{x^2}{2}\right\}$$

Интегральная предельная теорема Муавра-Лапласа.

Для нахождения значения вероятностей вида $P\{m_1 < m < m_2\}$ при заданных m_1 и m_2 и достаточно больших значениях n применяется другая формула, определяемая следующей теоремой.

Теорема. В условиях схемы Бернулли выполняется следующее равенство

$$\lim_{n\to\infty} P\left\{a < \frac{m-np}{\sqrt{np(1-p)}} < b\right\} = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\left\{-\frac{x^2}{2}\right\}dx = \Phi(b) - \Phi(a),$$

где $\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} \exp\left\{-\frac{x^2}{2}\right\}dx$ называется функцией Лапласа.

При достаточно больших значениях n можно записать

$$P\{m_1 < m < m_2\} \approx \Phi\left(\frac{m_2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{np(1-p)}}\right).$$

Теорема Пуассона. Распределение Пуассона

В условиях схемы Бернулли будем полагать, что n принимает достаточно большие значения, а значения вероятности p близки к нулю, то есть необходимо получить предельную формулу для $P_n(m)$ при выполнении двух предельных условий $n \to \infty$ и $p \to 0$.

Для того, чтобы сформулировать такую теорему, необходимо рассмотреть, так называемую **схему серий опытов**.

Пусть дана последовательность серий опытов: в первой серии единственный опыт и вероятность успеха – p_1; во второй серии два опыта и вероятность успеха в каждом из них – p_2; в третьей серии три опыта и вероятность успеха – p_3; в n-й серии n опытов и вероятность успеха в каждом из них – p_n.

Обозначим $np_n = a_n$. Будем полагать, что существует предел

$$\lim_{n\to\infty} np_n = \lim_{n\to\infty} a_n = a.$$

Теорема (Пуассона). В схеме серий, при выполнении условия

$$\lim_{n\to\infty} np_n = \lim_{n\to\infty} a_n = a,$$

выполняется равенство $\lim\limits_{n\to\infty} P_n(m) = P(m) = \dfrac{a^m}{m!}e^{-a}$.

Совокупность вероятностей $P(m)$ при всех $m = 0,1,2,\ldots$, называется **распределением Пуассона.**

Доказательство.

В силу равенства $np_n = a_n$, можно записать $p_n = a_n/n$, поэтому

$$P_n(m) = C_n^m p_n^m (1-p_n)^{n-m} = \frac{n!}{m!(n-m)!}(a_n/n)^m (1-a_n/n)^{n-m} =$$

$$= \frac{a_n^m}{m!}\frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m}(1-a_n/n)^{-m}(1-a_n/n)^n.$$

Так как выполняются равенства:

$$\lim_{n\to\infty}\frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m}=1,$$

$$\lim_{n\to\infty}\left(1-a_n/n\right)^{-m}=1,$$

$$\lim_{n\to\infty}\left(1-a_n/n\right)^{n}=e^{-a},$$

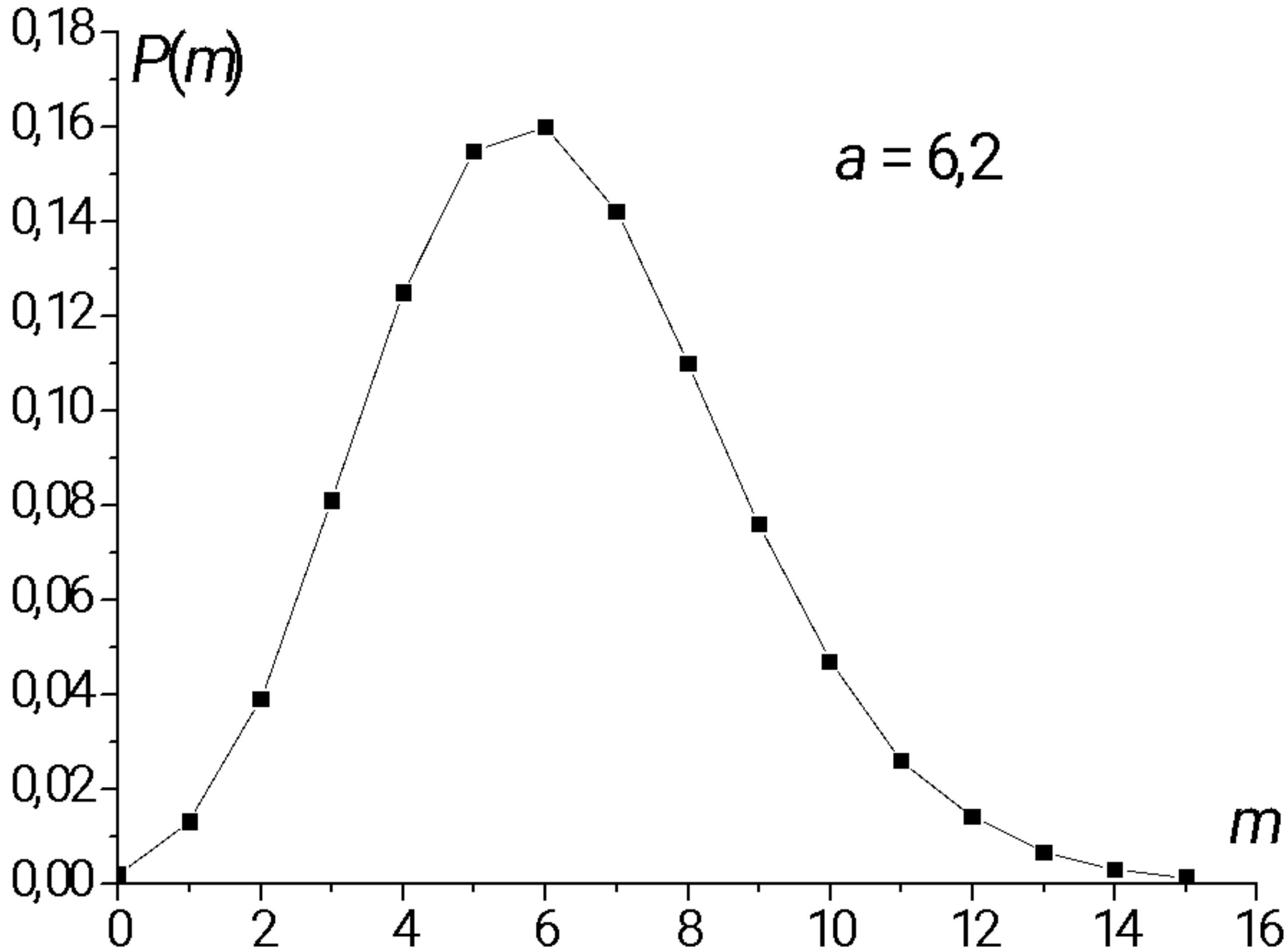
то можно записать

$$\lim_{n\to\infty}P_n(m)=P(m)=\frac{a^m}{m!}e^{-a}.$$

Рассмотрим вид графика $P(m)$ в зависимости от m. Для этого вычислим отношение

$$\frac{P(m)}{P(m-1)} = \frac{a^m e^{-a}}{m!} \cdot \frac{(m-1)!}{a^{m-1} e^{-a}} = \frac{a}{m}.$$

Отсюда видно, что при $m < a$ $P(m) > P(m-1)$, то есть $P(m)$ монотонно возрастает. Аналогично, при $m > a$ $P(m) < P(m-1)$, то есть $P(m)$ монотонно убывает. Своего максимального значения $P(m)$ достигает при $m =]a[$.



Вид графика $P(m)$ в зависимости от m при $a = 6{,}2$.

Случайные потоки однородных событий

Распределение Пуассона находит самое широкое применение в теории случайных потоков однородных событий.

Случайным потоком однородных событий называется ось времени, на которой в случайные моменты возникают некоторые однородные (одинаковые) события.

Отметим, что в теории потоков термин **событие** имеет смысл, существенно отличающийся от смысла понятия **случайного события** в теории вероятностей.

1. События наступают **независимо** друг от друга, то есть наступление какого-либо события не влияет на наступление других событий.

2. На бесконечно малом интервале времени продолжительности Δt событие может наступить с вероятностью $\lambda\Delta t + o(\Delta t)$, а два, и более событий наступают с вероятностью $o(\Delta t)$. Это свойство потоков называется **ординарностью** потока.

3. Параметр λ, который называется **интенсивностью потока**, не зависит от времени. Это свойство потоков называется **стационарностью потока.**

Поток, удовлетворяющий всем трём свойствам – независимости, ординарности и стационарности называется **пуассоновским потоком**.

Обозначим $P_t(m)$ вероятность того, что на интервале времени, продолжительности t наступит ровно m событий.

Для нахождения этой вероятности разделим рассматриваемый интервал времени на достаточно большое число n частей длительности t/n. Тогда в силу ординарности потока, при $n >> 1$, на каждом отрезке может наступить не более одного события, а каждый отрезок можно рассматривать как опыт, в котором может наступить или не наступить случайное событие – появление однородного события рассматриваемого потока. Вероятность этого случайного события составляет $p_n = \lambda t/n + o(n^{-1})$.

Тогда выполнены все условия теоремы Пуассона, поэтому для вероятности $P_t(m)$, в силу условия

$$\lim_{n\to\infty} np_n = \lim_{n\to\infty} n\left(\lambda t/n + o(n^{-1})\right) = \lambda t,$$

можно записать равенство

$$P_t(m) = \frac{(\lambda t)^m}{m!}\exp\{-\lambda t\},$$

которое определяет распределение Пуассона для числа событий, наступивших в пуассоновском потоке за время t.

Полиномиальное распределение

Обобщением схемы Бернулли, когда в каждом опыте возможно только два исхода – успех (наступление события A) или неудача (не наступление события A), рассматривается полиномиальная схема.

Аналогично схеме Бернулли (биномиальной схеме) в полиномиальной схеме также рассматривается серия из n независимых опытов, но в каждом из которых возможно наступление одного из K событий $A_1, A_2, \dots, A_K$, образующих полную группу попарно несовместных событий, с вероятностями $p_k = P(A_k)$ в одном опыте.

Требуется найти вероятность $P_n(m_1, m_2, \ldots, m_K)$ того, что в серии из n опытов событие A_1 наступит m_1 раз, событие A_2 наступит m_2 раза и так далее, наконец, событие A_K наступит m_K раз. Естественно, что выполняется равенство $m_1 + m_2 + \ldots + m_K = n$.

Проделав выкладки, аналогичные приведённым для схемы Бернулли, нетрудно показать, что выполняется равенство

$$P_n(m_1, m_2, \ldots, m_K) = \frac{n!}{m_1! m_2! \ldots m_K!} p_1^{m_1} p_2^{m_2} \ldots p_K^{m_K}.$$

Совокупность вероятностей $P_n(m_1, m_2, \ldots, m_K)$ для всех возможных целых, неотрицательных значений аргументов $m_1, m_2, \ldots, m_K$, удовлетворяющих условию $m_1 + m_2 + \ldots + m_K = n$, называется **полиномиальным распределением.**