Федеральное агентство по образованию
Государственное образовательное учреждение высшего профессионального образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

ЛЕКЦИЯ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

ТЕМА 7: СХЕМА БЕРНУЛЛИ

Лектор: Подберезина Е.И.
2010 год

Повторные испытания – это последовательное проведение n раз одного и того же опыта или одновременное проведение n одинаковых опытов. Например, при контроле уровня надежности прибора могут либо проводить n испытаний с одним и тем же прибором, если после отказа полностью восстанавливают его исходные свойства, либо ставить на испытания n опытных образцов этого прибора, которые считают идентичными.

Определение. ***Схемой Бернулли*** или ***последовательностью независимых одинаковых испытаний***, или ***биномиальной схемой испытаний*** называют последовательность n испытаний, удовлетворяющих условиям:

1) при каждом испытании различают два исхода: появление некоторого события A (успех) и появление противоположного ему события $\overline{A}$ (неудача);

2) испытания являются независимыми, то есть вероятность успеха в любом испытании не зависит от исходов всех предыдущих испытаний;

3) вероятность успеха в каждом испытании постоянна и равна $P(A) = p$; следовательно, вероятность неудачи во всех испытаниях тоже постоянна и равна $P(\overline{A}) = 1 - p = q$.

ПРИМЕР. Несколько последовательных бросаний монеты представляют собой независимые опыты. Несколько последовательных выниманий карты из колоды представляют собой независимые опыты при условии, что вынутая карта каждый раз возвращается в колоду и карты перемешиваются; в противном случае это – зависимые опыты. Несколько выстрелов представляют собой независимые опыты только в случае, если прицеливание производится заново перед каждым выстрелом; в случае, когда прицеливание производится один раз перед всей стрельбой или непрерывно осуществляется в процессе стрельбы (стрельба очередью, бомбометание серией), выстрелы представляют собой зависимые опыты.

Приведем примеры испытаний по схеме Бернулли.

ПРИМЕР. Последовательное подбрасывание n раз симметричной и однородной монеты. Здесь успехом является появление герба. Вероятность успеха равна $p = 0{,}5$; вероятность неудачи $q = 0{,}5$.

ПРИМЕР. Последовательное бросание n раз игральной кости. Здесь успехом можно считать, например, появление шестерки. Вероятность успеха равна $p = \frac{1}{6}$, вероятность неудачи $q = \frac{5}{6}$.

ПРИМЕР. Последовательность n выстрелов стрелка по мишени можно лишь приближенно рассматривать как схему испытаний Бернулли, так как независимость результатов стрельбы может нарушаться либо из-за «пристрелки» спортсмена, либо вследствие его утомляемости.

При рассмотрении испытаний по схеме Бернулли основной задачей является нахождение вероятности события A_k = < в n испытаниях успех произойдет ровно k и, значит, не произойдет ровно $n - k$ раз >. Вероятность события A_k обозначают $P(A_k) = P_n(k)$ и вычисляют по следующей теореме.

ТЕОРЕМА. Вероятность $P_n(k)$ того, что в n испытаниях по схеме Бернулли произойдет ровно k успехов (ровно k раз произойдет событие A) равна

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}, \quad k = 0, 1, \ldots, n.$$

Последняя формула называется ***формулой Бернулли***.

ДОКАЗАТЕЛЬСТВО. Существует ровно C_n^k разных способов выбрать k номеров испытаний, в которых наступит событие A. Таким образом, существует C_n^k несовместных событий, каждое из которых представляет собой произведение k событий A и $n - k$ событий $\overline{A}$. Так как события A и $\overline{A}$ независимы, то вероятность любого из этих несовместных событий равна $p^k \cdot q^{n-k}$. Сумма всех этих C_n^k несовместных событий равна событию A_k =< в n испытаниях успех произойдет ровно k и, значит, не произойдет ровно $n - k$ раз >. По теореме сложения имеем:

$$P(A_k) = P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}, \quad k = 0, 1, \ldots, n.$$

Теорема доказана.

Если число испытаний n велико, использование формулы Бернулли связано с необходимостью выполнять громоздкие вычисления. В этом случае используют приближенные формулы, которые называются ***асимптотическими*** и определяются теоремой Пуассона, локальной и интегральной теоремами Муавра-Лапласа. Эти формулы тем точнее, чем больше число испытаний.

Когда n имеет порядок не менее нескольких десятков, а лучше нескольких сотен, и произведение $n \cdot p > 10$, пользуются локальной теоремой Муавра-Лапласа. Для $p = 0{,}5$ она была получена в 1783 году Муавром, в 1783 году Лаплас обобщил формулу Муавра на случай произвольного p, $p \neq 0$, $p \neq 1$.

ТЕОРЕМА. (***Локальная теорема Муавра-Лапласа***). Если вероятность p наступления события A в каждом испытании постоянна и отлична от 0 и 1, то вероятность $P_n(k)$ того, что в n независимых испытаниях событие A произойдет ровно k раз, приближенно равна

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \qquad (9)$$

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ – ***функция Гаусса*** и $x = \frac{k - np}{\sqrt{npq}}$.

Чем больше n, тем точнее приближенная формула (9), называемая ***локальной формулой Муавра-Лапласа***. Приближенные значения вероятности $P_n(k)$, получаемые по этой формуле, на практике используются как точные при npq порядка двух и более десятков, то есть при условии $npq \geq 20$.

ЗАМЕЧАНИЕ. Для упрощения расчетов, связанных с применением формулы (9), составлена таблица значений функции $\varphi(x)$. Пользуясь этой таблицей, необходимо знать следующие свойства функции $\varphi(x)$: функция $\varphi(x)$ является четной: $\varphi(x) = \varphi(-x)$; функция $\varphi(x)$ монотонно убывает при положительных значениях x, причем при $x \to \infty$ функция $\varphi(x) \to 0$. Практически можно считать, что $\varphi(x) \approx 0$ уже при $x > 4$.

ПРИМЕР. Два спортсмена играют в настольный теннис. Вероятность выигрыша первого спортсмена равна $\frac{5}{9}$. Какова вероятность того, что он выиграет две партии из пяти?

Решение. По условию имеем: A = < первый спортсмен выиграл партию >; $p = \frac{5}{9}$, $q = \frac{4}{9}$, $n = 5$, $k = 2$. Требуется найти вероятность $P_5(2)$ двух успехов в пяти испытаниях. Воспользуемся локальной формулой Муавра-Лапласа.

$$x = \frac{k - np}{\sqrt{npq}} = \frac{2 - 5 \cdot \frac{5}{9}}{\frac{10}{9}} = -0{,}7.$$

По таблице находим $\varphi(-0{,}7) = \varphi(0{,}7) = 0{,}3123$. По формуле (9) имеем:

$$P_5(2) = \frac{9}{10} \cdot \varphi(-0{,}7) = \frac{9}{10} \cdot 0{,}3123 \approx 0{,}281.$$

Проверим результат, воспользовавшись формулой Бернулли:

$$P_5(2) = C_5^2 \left(\frac{5}{9}\right)^2 \left(\frac{4}{9}\right)^3 \approx 0{,}271.$$

Такое расхождение объясняется тем, что локальная теорема Муавра-Лапласа дает хорошее приближение при большом n, а здесь $n = 5$.

Если вероятность p события A в одном испытании близка к нулю (такие события называются редкими) n велико и $n\,p < 10$, то вероятность $P_n(k)$ вычисляют по формуле Пуассона.

ТЕОРЕМА. Если вероятность p наступления события A в каждом испытании постоянна, близка к нулю, число независимых испытаний n велико, а произведение $n\,p = \lambda < 10$, то вероятность $P_n(k)$ того, что в n независимых испытаниях событие A произойдет ровно k раз, приближенно равна

$$P_n(k) \approx \frac{\lambda^k \cdot e^{-\lambda}}{k!}. \tag{10}$$

Формула (10) называется ***формулой Пуассона***.

ПРИМЕР. Некоторое электронное устройство выходит из строя, если откажет определенная микросхема. Вероятность ее отказа в течение одного часа работы устройства равна 0,004. Какова вероятность того, что за 1000 часов работы устройства придется 5 раз менять микросхему?

Решение. Успехом является событие $A = <$ отказ микросхемы в течение часа работы электронного устройства $>$, $n = 1000$, $p = 0{,}004$. Так как $\lambda = n\,p = 4 < 10$, то следует использовать формулу Пуассона:

$$P_{1000}(5) \approx \frac{4^5 \cdot e^{-4}}{5!} \approx 0{,}1563.$$

Найдем вероятность того же события по теореме Муавра-Лапласа:

$$x = \frac{k - n\,p}{\sqrt{npq}} = \frac{5-4}{1{,}996} \approx 0{,}501.$$

$\varphi(0{,}501) \approx 0{,}3521$. Тогда

$$P_{1000}(5) \approx \frac{1}{\sqrt{n\,p\,q}} \cdot \varphi(x) \approx \frac{0{,}3521}{1{,}996} \approx 0{,}1764.$$

Значение искомой вероятности по формуле Бернулли равно

$$P_{1000}(5) = C_{1000}^{5}\,(0{,}004)^5\,(0{,}996)^{995} \approx 0{,}1566.$$

Как видно из полученных результатов, формула Пуассона дает в данном случае более точный результат, чем локальная формула Муавра-Лапласа.

ТЕОРЕМА. (***Интегральная теорема Муавра-Лапласа***). Если вероятность p наступления события A в каждом испытании постоянна и отлична от 0 и 1, то вероятность $P_n(k_1, k_2)$ того, что число наступлений события A в n независимых испытаниях заключено в пределах от k_1 до k_2 включительно, при достаточно большом числе n испытаний приближенно равна

$$P_n(k_1, k_2) \approx \Phi(x_2) - \Phi(x_1), \qquad (11)$$

где $\Phi(x) = \int\limits_0^x \varphi(y)\,dy = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{y^2}{2}}\,dy$ – ***интеграл Лапласа***,

$$x_1 = \frac{k_1 - n\,p}{\sqrt{n\,pq}}, \quad x_2 = \frac{k_2 - n\,p}{\sqrt{n\,pq}}.$$

Чем больше n, тем точнее формула (11), которая называется ***интегральной формулой Муавра-Лапласа***. Интегральная формула Муавра-Лапласа, так же как и локальная формула Муавра-Лапласа, дает удовлетворительную для практики погрешность вычисления вероятностей при выполнении условия $n\,p\,q \geq 20$.

ЗАМЕЧАНИЕ. Для $0 \leq x \leq 5$ составлена таблица значений функции $\Phi(x)$. Для ее использования нужно знать свойства функции $\Phi(x)$. Легко показать, что эта функция нечетная: $\Phi(-x) = -\,\Phi(x)$. Функция $\Phi(x)$ монотонно возрастает, причем при $x \to +\infty$ функция $\Phi(x) \to 1$. Практически можно считать, что $\Phi(x) \approx 1$ уже при $x > 4$.

Сформулируем важное следствие теоремы Муавра-Лапласа.

СЛЕДСТВИЕ. Вероятность того, что в n независимых испытаниях, в каждом из которых вероятность появления события A равна p, $p \neq 0$, $p \neq 1$, абсолютная величина отклонения относительной частоты появления события A от вероятности появления этого события не превышает положительного числа ε, приближенно равна удвоенному интегралу Лапласа при $x = \varepsilon \cdot \sqrt{\frac{n}{pq}}$:

$$P\left(\left| \frac{m}{n} - p \right| \leq \varepsilon \right) = 2 \cdot \Phi\left(\varepsilon \cdot \sqrt{\frac{n}{pq}} \right).$$

ПРИМЕР. Вероятность изготовления на автоматическом станке стандартной детали равна 0,8. Найти вероятности возможного числа появления бракованных деталей среди пяти отобранных.

Решение. Успехом является событие A = < изготовлена бракованная деталь >. Имеем: $n = 5$, $p = 1 - 0{,}8 = 0{,}2$; $q = 0{,}8$. По формуле Бернулли находим:

$$P_5(0) = C_5^0 \cdot (0{,}2)^0 \cdot (0{,}8)^5 = 0{,}32768;$$
$$P_5(1) = C_5^1 \cdot (0{,}2)^1 \cdot (0{,}8)^4 = 0{,}4096;$$
$$P_5(2) = C_5^2 \cdot (0{,}2)^2 \cdot (0{,}8)^3 = 0{,}2048;$$
$$P_5(3) = C_5^3 \cdot (0{,}2)^3 \cdot (0{,}8)^2 = 0{,}0512;$$
$$P_5(4) = C_5^4 \cdot (0{,}2)^4 \cdot (0{,}8)^1 = 0{,}0064;$$
$$P_5(5) = C_5^5 \cdot (0{,}2)^5 \cdot (0{,}8)^0 = 0{,}00032.$$

Мы видим, что число $k_0 = 1$ бракованных изделий среди 5 отобранных обладает наибольшей вероятностью.

Определение. Число k_0 наступлений события A называется ***наивероятнейшим***, если оно имеет наибольшую вероятность $P_n(k_0)$ по сравнению с вероятностями $P_n(k)$ наступления события A любое другое количество раз.

Наивероятнейшее число k_0 находят из двойного неравенства:

$$np - q \le k_0 \le np + p,$$

причем: а) если число $np - q$ дробное, то число k_0 единственное; б) если число $np - q$ целое, то существует два наивероятнейших числа: $k_0 = np - q$ и $k_0 = np + p$; в) если число np – целое, то наивероятнейшее число $k_0 = np$.