11.Условная вероятность. Независимость событий.

Начнем с примера. Рассмотрим задачу с однократным подбрасыванием симметричного кубика. Введем события: A - выпадение двух очков, B - выпадение четного числа очков. Вероятности каждого из этих событий легко рассчитать по классической схеме:

$$P(A) = \frac{1}{6}, \qquad P(B) = \frac{3}{6} = \frac{1}{2}.$$

Зададим себе следующий вопрос: как изменится вероятность события A, если стало известно, что произошло событие B, т.е. выпало четное число очков ? Применим снова классическую схему. Число возможных исходов уменьшилось - их стало 3: выпадение двойки, четверки, шестерки. Поэтому искомая вероятность равна 1/3.

Перейдем к общей классической схеме. Пусть пространство Ω содержит n равновозможных элементарных событий. Рассмотрим события A и B с m и k благоприятствующими исходами соответственно. По классической схеме

$$P(A) = \frac{m}{n}, \qquad P(B) = \frac{k}{n}.$$

Найдем вероятность события A при условии, что событие B произошло. Будем обозначать ее далее как $P(A \mid B)$. Тогда, по классической схеме ее следует определить как

$$P(A \mid B) = \frac{r}{k},$$

где r - число элементарных событий, благоприятствующих и A и B одновременно, т.е. благоприятствующих событию AB. Но тогда

$$P(A \mid B) = \frac{r}{k} = \frac{r/n}{k/n} = \frac{P(AB)}{P(B)}.$$

Перейдем теперь к общему определению.

Определение.

Пусть задано вероятностное пространство $\langle \Omega, \Im, P \rangle$ и события $A, B \in \Im$. Если $P(B) > 0$, то вероятность события A при условии, что произошло событие B, называемая *условной вероятностью*, полагается равной

$$P(A \mid B) = \frac{P(AB)}{P(B)}.$$

Пусть событие B с $P(B) > 0$ фиксировано. Легко проверить, что функция $P_B(A) = P(AB)/P(B)$, определенная для всех $A \in \Im$, удовлетворяет аксиомам $A1 - A3$. Значит, для нее справедливы и все следствия, вытекающие из аксиом.

Введенное понятие условной вероятности позволяет естественным образом определить независимость событий. В самом деле, равенство

$$P(A \mid B) = P(A),$$

показывающее, что вероятность события A никак не изменяется при наступлении события B, согласуется с нашим интуитивным представлением о независимости событий. Однако, за определение независимости берут симметричное условие, которое получается домножением на $P(B) > 0$.

Определение. События A и B $(A, B \in \Im)$ называются *независимыми*, если

$$P(AB) = P(A)P(B).$$

Простейшим примером независимых событий может служить выпадение герба при первом бросании монеты и выпадение решки при втором бросании. Однако, можно указать такие события и в эксперименте с однократным подбрасыванием. Так, при однократном бросании игральной кости такими событиями будут: A - выпадение четного числа очков, B - выпадение числа очков, кратного трем. Действительно,

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{3}, \quad P(AB) = P(\omega_6) = \frac{1}{6}.$$

Пусть имеется три попарно независимых события A, B, C, т.е.

$$P(AB) = P(A)P(B), \quad P(BC) = P(B)P(C), \quad P(AC) = P(A)P(C).$$

Возникает вопрос: будет ли при этом справедлива следующая формула

$$P(ABC) = P(A)P(B)P(C) \quad ?$$

Ответ на него заключен в известном примере Бернштейна.

Пример Бернштейна. Рассмотрим следующий эксперимент. На плоскость бросается тетраэдр, три грани которого выкрашены соответственно в красный, синий и зеленый цвета, а на четвертую нанесены все три цвета (триколор). Пусть R - событие, состоящее в том, что тетраэдр упал на грань, содержащую красный цвет; B - упал на грань, содержащую синий цвет; G - упал на грань, содержащую зеленый цвет. Определим вероятности каждого из введенных событий по классической схеме:

$$P(R) = \frac{2}{4} = \frac{1}{2}, \quad P(B) = \frac{1}{2}, \quad P(G) = \frac{1}{2}.$$

Вычислим теперь вероятности попарных произведений событий:

$$P(RB) = \frac{1}{4}, \quad P(RG) = \frac{1}{4}, \quad P(BG) = \frac{1}{4}$$

- поскольку два цвета одновременно содержит лишь одна грань (триколор).

Отсюда следует попарная независимость событий:

$$P(RB) = P(R)P(B), \quad P(RG) = P(R)P(G), \quad P(BG) = P(B)P(G).$$

В то же время,

$$P(RBG) = \frac{1}{4}, \text{ а } P(R)P(B)P(G) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8},$$

и, следовательно,

$$P(RBG) \neq P(R)P(B)P(G).$$

Введем новое определение, касающееся независимости событий.

Определение. События $A_1, A_2, \ldots, A_n \in \Im$ называются *взаимно независимыми* или *независимыми в совокупности*, если для $\forall i_1, i_2, \ldots, i_k$ таких, что $1 \leq i_1 < i_2 < \ldots < i_k \leq n \quad (k = 2, 3, \ldots, n)$,

$$P\left(A_{i_1} A_{i_2} \ldots A_{i_k}\right) = P\left(A_{i_1}\right) P\left(A_{i_2}\right) \ldots P\left(A_{i_k}\right).$$

Из взаимной независимости событий следует их попарная независимость. Обратное неверно, что следует из рассмотренного примера Бернштейна.

12.Теоремы сложения и умножения вероятностей.

В следствиях, вытекающих из аксиом теории вероятностей, была доказана формула для вероятности суммы двух произвольных событий A и B:

$$P\left(A \cup B\right) = P\left(A\right) + P\left(B\right) - P\left(AB\right).$$

Попытаемся теперь установить формулу для вероятности суммы произвольного числа событий. Начнем с суммы трех событий A, B, C. Вводя временное переобозначение $D = B \cup C$, в соответствии с формулой для суммы двух событий имеем последовательность равенств:

$$P\left(A \cup B \cup C\right) = P\left(A \cup D\right) = P\left(A\right) + P\left(D\right) - P\left(AD\right) =$$

$$= P\left(A\right) + P\left(B \cup C\right) - P\left(A\left(B \cup C\right)\right) =$$

$$= P\left(A\right) + P\left(B\right) + P\left(C\right) - P\left(BC\right) - P\left(AB \cup AC\right) =$$

$$= P\left(A\right) + P\left(B\right) + P\left(C\right) - P\left(BC\right) - P\left(AB\right) - P\left(AC\right) + P\left(ABAC\right).$$

Учитывая, что $A^2 = A$, окончательно находим:

$$P\left(A \cup B \cup C\right) = P\left(A\right) + P\left(B\right) + P\left(C\right) - P\left(AB\right) - P\left(AC\right) - P\left(BC\right) + P\left(ABC\right).$$

По аналогии с полученным соотношением легко записать формулу для вероятности суммы n событий $A_1, A_2, \ldots, A_n$:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P\left(A_i\right) - \sum_{1 \leq i} \sum_{< j \leq n} P(A_i A_j) +$$

$$+ \sum_{1 \leq i} \sum_{< j} \sum_{< k \leq n} P(A_i A_j A_k) + \ldots + (-1)^{n-1} P(\bigcap_{k=1}^{n} A_k).$$

Эту сложную формулу называют *теоремой сложения* вероятностей. Она доказывается методом математической индукции. В случае попарно несовместных событий: $A_i A_j = \emptyset$ при $i \neq j$ формула приобретает вид аксиомы конечной аддитивности:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P\left(A_i\right).$$

Формула для вероятности произведения двух событий следует из определения условной вероятности:

$$P(AB) = P(A)\,P(B \mid A).$$

Для произведения трех событий A, B, C, вводя временное переобозначение $D = AB$, имеем:

$$P(ABC) = P(DC) = P(D)\,P(C \mid D) = P(AB)\,P(C \mid AB) =$$
$$= P(A)\,P(B \mid A)P(C \mid AB).$$

Итак,

$$P(ABC) = P(A)\,P(B \mid A)P(C \mid AB).$$

Для произведения n событий $A_1, A_2, \ldots, A_n$ по аналогии можно записать:

$$P(\bigcap_{i=1}^{n} A_i) = P(A_1)\,P(A_2 \mid A_1)P(A_3 \mid A_1A_2) \cdot \ldots \cdot P(A_n \mid A_1 \ldots A_{n-1}).$$

Эта формула называется *теоремой умножения* вероятностей. Докажем ее методом математической индукции. Рассмотрим вероятность произведения $(n+1)$ события

$$P(\bigcap_{i=1}^{n+1} A_i) = P(A_{n+1} \cdot \bigcap_{i=1}^{n} A_i).$$

Согласно формулам условной вероятности и вероятности произведения n событий

$$P(\bigcap_{i=1}^{n+1} A_i) = P(\bigcap_{i=1}^{n} A_i)P(A_{n+1} \mid \bigcap_{i=1}^{n} A_i) =$$

$$= P(A_1)\,P(A_2 \mid A_1) \cdot \ldots \cdot P(A_n \mid A_1 \ldots A_{n-1})P(A_{n+1} \mid A_1 \ldots A_{n-1}A_n),$$

что и доказывает формулу умножения вероятностей.

Заметим, что для взаимно независимых событий $A_1, A_2, \ldots, A_n$ все условные вероятности превращаются в безусловные, и формула умножения переходит в

$$P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)\,.$$

Пример. Задача о неорганизованных зрителях.

В зрительном зале n пронумерованных мест, и все билеты на вечерний спектакль проданы. Однако зрители, придя в театр, рассаживаются не согласно купленным билетам, а случайным образом. Найти вероятность того, что ни один из зрителей не сядет на "свое" (указанное в билете) место.

Введем следующие события: A_i - i-тый зритель садится на "свое" место $\left(i = \overline{1, n}\right)$; B - ни один из зрителей не сядет на "свое" место. Будем искать вероятность противоположного B события $\bar{B}$, состоящего в том, что хотя бы один из зрителей сядет на "свое" место. Это событие является суммой введенных событий A_i

$$\bar{B} = A_1 \cup A_2 \cup \ldots \cup A_n.$$

Применяя теорему сложения вероятностей, найдем:

$$P\left(\bar{B}\right)=\sum_{k=1}^{n}\left(-1\right)^{k-1}\sum_{1\le i_1<}\sum_{i_2<\ldots}\ldots\sum_{<i_k\le n}P(A_{i_1}A_{i_2}\cdot\ldots\cdot A_{i_k}).$$

Вычислим вероятности произведения событий по теореме умножения вероятностей, применяя классическую схему:

$$P(A_{i_1})=\frac{1}{n}\text{ для }\forall i_1;$$

$$P(A_{i_1}A_{i_2})=P(A_{i_1})P(A_{i_2}\mid A_{i_1})=\frac{1}{n}\cdot\frac{1}{n-1}\text{ для }\forall i_1,i_2;$$

и т.д.

$$P(A_{i_1}A_{i_2}\cdot\ldots\cdot A_{i_k})=P(A_{i_1})P(A_{i_2}\mid A_{i_1})\cdot\ldots\cdot P(A_{i_k}\mid A_{i_1}\cdot\ldots\cdot A_{i_{k-1}})=$$

$$=\frac{1}{n}\cdot\frac{1}{n-1}\cdot\ldots\cdot\frac{1}{n-(k-1)}=\frac{(n-k)!}{n!}\text{ для }\forall i_1,i_2,\ldots,i_k.$$

Подставляя найденные вероятности в выражение для $P\left(\bar{B}\right)$, получаем:

$$P\left(\bar{B}\right)=\sum_{k=1}^{n}\left(-1\right)^{k-1}\frac{(n-k)!}{n!}\sum_{1\le i_1<}\sum_{i_2<\ldots}\ldots\sum_{<i_k\le n}1.$$

Число слагаемых в сумме равно числу выборок без возвращения k номеров из $(1,2,\ldots,n)$, отличающихся составом, т.е. совпадает с числом сочетаний C_n^k. Поэтому

$$P\left(\bar{B}\right)=\sum_{k=1}^{n}\left(-1\right)^{k-1}\frac{(n-k)!}{n!}C_n^k=\sum_{k=1}^{n}\frac{\left(-1\right)^{k-1}}{k!}.$$

Таким образом, вероятность искомого события

$$P\left(B\right)=1-P\left(\bar{B}\right)=1-\sum_{k=1}^{n}\frac{\left(-1\right)^{k-1}}{k!}=\sum_{k=0}^{n}\frac{\left(-1\right)^{k}}{k!}.$$

Заметим, что полученный ряд является быстро сходящимся, и при больших n найденная вероятность приблизительно равна

$$P\left(B\right)=\sum_{k=0}^{\infty}\frac{\left(-1\right)^{k}}{k!}=\frac{1}{e}\approx 0,37.$$

13.Формула полной вероятности и формула Байеса.

Определение. Говорят, что события $\{H_1,H_2,\ldots,H_n\}$ образуют *полную группу* попарно несовместных событий, если:

1) $H_1\cup H_2\cup\ldots\cup H_n=\Omega$,

2) $H_iH_j=\emptyset$ при $i\ne j$.

Теорема. Пусть A - произвольное событие из алгебры $\Im$ и $H_1, H_2, \ldots, H_n$ ($H_i \in \Im$) - группа попарно несовместных событий: $H_i H_j = \emptyset$ при $i \neq j$; причем $A \subset H_1 \cup H_2 \cup \ldots \cup H_n$. Тогда имеет место формула ($P(H_k) > 0$)

$$P(A) = \sum_{k=1}^{n} P(H_k)\, P(A \mid H_k),$$

называемая *формулой полной вероятности*.

Доказательство: из условия $A \subset H_1 \cup H_2 \cup \ldots \cup H_n$ следует, что

$$A = A(H_1 \cup H_2 \cup \ldots \cup H_n) = AH_1 \cup AH_2 \cup \ldots \cup AH_n.$$

Отсюда по аксиоме конечной аддитивности для несовместных событий и определению условной вероятности

$$P(A) = P(\bigcup_{k=1}^{n} AH_k) = \sum_{k=1}^{n} P(AH_k) = \sum_{k=1}^{n} P(H_k)\, P(A \mid H_k). \qquad \boxplus$$

Замечание. Поскольку для $\forall A \in \Im : A \subset \Omega$ формула полной вероятности справедлива, в частности, для полной группы попарно несовместных событий.

Пример. Задача о стратегии сдачи экзамена.

Студент пришел на экзамен, выучив m билетов из n. В каком случае его шансы сдать экзамен выше: когда он приходит на экзамен первым или когда приходит вторым ?

Введем события A_1 - студент сдаст экзамен, придя на него первым; A_2 - студент сдаст экзамен, придя на него вторым. Совершенно ясно, что

$$P(A_1) = \frac{m}{n}.$$

Для отыскания вероятности $P(A_2)$ необходимо знать, какой билет вытащил студент, пришедший на экзамен первым. Поэтому рассмотрим полную группу попарно несовместных событий: H - студент, пришедший на экзамен первым, вытащил один из "хороших" билетов, $\bar{H}$ - он вытащил один из "плохих" билетов. Расписывая $P(A_2)$ по формуле полной вероятности, находим:

$$P(A_2) = P(H)\, P(A_2 \mid H) + P\left(\bar{H}\right) P(A_2 \mid \bar{H}) =$$

$$= \frac{m}{n} \cdot \frac{m-1}{n-1} + \frac{n-m}{n} \cdot \frac{m}{n-1} = \frac{m\,(m-1+n-m)}{n\,(n-1)} = \frac{m}{n}.$$

Получился довольно неожиданный ответ: шансы сдать экзамен оказались одинаковыми. Более того, если бы этот студент пришел на экзамен k-ым, вероятность сдачи им экзамена осталась бы точно такой же.

Рассмотрим опять группу попарно несовместных событий $\{H_1, H_2, \ldots, H_n\}$, которые трактуются в качестве *гипотез* о причинах возникновения какого-то явления. Далее ставится эксперимент, в результате которого происходит некоторое событие

$A \subset H_1 \cup H_2 \cup \ldots \cup H_n$. Поинтересуемся, как изменились вероятности гипотез после наступления события A, т.е. вычислим условные вероятности $P(H_i \mid A)$.

Согласно определению условной вероятности

$$P(H_i \mid A) = \frac{P(AH_i)}{P(A)} = \frac{P(H_i)P(A \mid H_i)}{P(A)}.$$

Применяя для $P(A)$ формулу полной вероятности, окончательно находим:

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{\sum\limits_{k=1}^{n} P(H_k)\, P(A \mid H_k)}.$$

Эта формула была установлена в 1750 г. Томасом Байесом - учеником де Муавра, поэтому ее называют *формулой Байеса*. Формула Байеса связывает апостериорные (послеопытные) вероятности гипотез $P(H_i \mid A)$ с априорными (доопытными) вероятностями $P(H_i)$. Она, как следует из вывода, является, фактически, другой формой записи формулы полной вероятности и формулы для условной вероятности.

Формула Байеса представляет собой теорему о вероятностях причин. В большинстве применений априорные вероятности $P(H_i)$ неизвестны, поэтому их для простоты полагают одинаковыми. Однако, такой подход не всегда приемлем. Незнание априорных вероятностей оказалось столь разрушительным для обоснованности статистических выводов из теоремы Байеса, что эта теорема была практически исключена из статистических исследований. Однако, в середине XX века байесовский подход получил развитие. Стали после каждого наблюдения апостериорные вероятности пересчитывать и на следующем шаге использовать как априорные. Это снизило эффект априорной неопределенности.

Пример. Задача о двух студентах.

Преподаватель математики дал двум студентам большое количество стандартных задач и попросил за академический час решить их как можно больше. Известно, что первый студент (назовем его условно "Голова") решает в k раз больше задач по сравнению со вторым студентом ("Тормоз") за один и тот же промежуток времени. После истечения академического часа преподаватель собрал решенные задачи, перемешал их и, выбрав наугад одну, проверил. Задача оказалась решенной неправильно. Какова вероятность того, что эту задачу решал первый студент, если процент "брака" у него P_1, а процент "брака" у второго студента - P_2 ?

Введем две очевидные гипотезы: H_1 - задачу решал первый студент, H_2 - задачу решал второй студент и событие A - выбранная задача решена неправильно. Нам необходимо найти условную вероятность $P(H_1 \mid A)$.

Поскольку $H_1 H_2 = \emptyset$ и $A \subset H_1 \cup H_2$ справедлива формула Байеса

$$P(H_1 \mid A) = \frac{P(H_1)P(A \mid H_1)}{P(H_1)\, P(A \mid H_1) + P(H_2)\, P(A \mid H_2)}.$$

Найдем априорные вероятности $P(H_1)$ и $P(H_2)$. Допустим, что "Тормоз" решил за академический час m задач. Тогда "Голова" решил $m \cdot k$ задач. В результате, согласно классической схеме

$$P(H_1) = \frac{m \cdot k}{m \cdot k + m} = \frac{k}{k+1}, \quad P(H_2) = \frac{1}{k+1}.$$

Учитывая, что условные вероятности $P(A \mid H_1) = P_1$, $P(A \mid H_2) = P_2$, имеем:

$$P(H_1 \mid A) = \frac{kP_1}{kP_1 + P_2}.$$

14.Схема независимых испытаний Бернулли.

Представим себе, что последовательно проводятся n испытаний, в каждом из которых интересующее нас событие A может произойти, а может и не произойти. Опыты повторяются в одних и тех же условиях, поэтому результат предыдущего опыта не влияет на последующий. Будем также предполагать, что вероятность появления события A не зависит от номера испытания: $P(A) = p$. Такую последовательность независимых испытаний называют *схемой Бернулли.*

Определим вероятность события B_k - появления события A ровно k раз в n испытаниях. Наступление события A будем называть "успехом", а его непоявление - "неудачей".

Теорема. Для вероятности k успехов в n испытаниях Бернулли справедлива формула

$$P(B_k) = P_{k;n} = C_n^k p^k q^{n-k},$$

где $k = 0, 1, 2, \ldots, n; \quad q = 1 - p.$

Доказательство:

Введем вспомогательные события S_i - успех в i-том опыте, F_j - неудача в j-том испытании. При этом $P(S_i) = P(A) = p$. Поскольку $F_j = \bar{S}_j$, то $P(F_j) = P(\bar{A}) = 1 - p = q$. Элементарными событиями в данной схеме являются цепочки вида

$$\omega = S_1 S_2 F_3 \ldots F_{n-2} S_{n-1} F_n.$$

Каждая благоприятная элементарная цепочка содержит ровно k успехов. Выразим событие B_k через них:

$$B_k = \bigcup_{(i_1, i_2, \ldots, i_k) \subset (1, 2, \ldots, n)} S_{i_1} S_{i_2} \ldots S_{i_k} F_{i_{k+1}} \ldots F_{i_n}.$$

Количество слагаемых в сумме совпадает с числом неупорядоченных выборок k различных чисел из $(1, 2, \ldots, n)$, т.е. с числом сочетаний C_n^k.

Согласно аксиоме конечной аддитивности

$$P(B_k) = \sum_{(i_1, i_2, \ldots, i_k) \subset (1, 2, \ldots, n)} P(S_{i_1} S_{i_2} \ldots S_{i_k} F_{i_{k+1}} \ldots F_{i_n}) =$$

и, в силу независимости событий, относящихся к разным испытаниям,

$$= \sum_{(i_1, i_2, \ldots, i_k) \subset (1, 2, \ldots, n)} P\left(S_{i_1}\right) P\left(S_{i_2}\right) \ldots P\left(S_{i_k}\right) P\left(F_{i_{k+1}}\right) \ldots P\left(F_{i_n}\right) =$$

$$= \sum_{(i_1, i_2, \ldots, i_k) \subset (1, 2, \ldots, n)} p^k q^{n-k} = C_n^k p^k q^{n-k}.$$

Итак, окончательно

$$P_{k;n} = P(B_k) = C_n^k p^k q^{n-k}. \qquad \boxtimes$$

Такое распределение получило название *биномиального закона.*

Замечая, что

$$\bigcup_{k=0}^{n} B_k = \Omega,$$

находим:

$$P(\bigcup_{k=0}^{n} B_k) = \sum_{k=0}^{n} P(B_k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1 = P(\Omega).$$

Отсюда и проистекает название закона.

Пример. Как играть в проигрышную игру.

Рассмотрим некоторую игру с четным числом партий $n = 2k$, причем каждая партия всегда заканчивается победой одного из участников. Пусть игрок A играет хуже игрока B, т.е. он выигрывает отдельную партию с вероятностью $p < 1/2$, а проигрывает с вероятностью $q = 1 - p > 1/2$. Подобная игра для игрока A считается проигрышной. Допустим, что игрок A имеет возможность выбирать количество партий n. Как ему поступить, если для общей победы любому из участников нужно выиграть больше половины партий ?

На первый взгляд, кажется очевидным следование общему принципу: "Чем раньше мы прекратим проигрышную игру, тем лучше для нас". Поэтому игроку A следует выбрать минимально возможное количество партий $n = 2$. Однако, подобная стратегия не всегда является оптимальной.

В самом деле. Рассматриваемая задача подпадает под схему независимых испытаний Бернулли, поэтому вероятность общей победы игрока A равна

$$P_A = P_{k+1;2k} + P_{k+2;2k} + \ldots + P_{2k;2k} = \sum_{i=k+1}^{2k} C_{2k}^i p^i q^{2k-i}.$$

Оптимальное количество партий n нужно выбирать из условия максимума вероятности P_A, которая является функцией дискретного параметра k и непрерывного параметра $p \quad (0 < p < 1/2) : P_A = W_{2k}(p)$. Поскольку число играемых партий четно, необходимо сравнить вероятности выигрыша для $2k-2$ и $2k$ партий, т.е. решить неравенство

$$\frac{W_{2k}(p)}{W_{2k-2}(p)} > 1$$

или

$$W_{2k}(p) - W_{2k-2}(p) > 0.$$

Запишем выражение для $W_{2k}(p)$ в интегральной форме. Дифференцируя $W_{2k}(p)$ по p, найдем:

$$W'_{2k}(p) = \sum_{i=k+1}^{2k} C_{2k}^i i p^{i-1} q^{2k-i} - \sum_{i=k+1}^{2k-1} C_{2k}^i (2k-i) p^i q^{2k-i-1} =$$

$$= 2k \cdot \left(\sum_{i=k+1}^{2k} C_{2k-1}^{i-1} p^{i-1} q^{2k-i} - \sum_{i=k+1}^{2k-1} C_{2k-1}^{i} p^{i} q^{2k-i-1} \right).$$

Меняя в первой сумме индекс суммирования $i-1=m$, получаем простой результат:

$$W'_{2k}(p) = 2kC_{2k-1}^{k} p^k q^{k-1} = kC_{2k}^{k} p^k q^{k-1}.$$

Интегрируя по p это соотношение с учетом того, что $W_{2k}(0)=0$, имеем:

$$W_{2k}(p) = kC_{2k}^{k} \int\limits_0^p t^k (1-t)^{k-1}\, dt.$$

Подставляя это выражение в неравенство и объединяя интегралы, придем к

$$\int\limits_0^p t^{k-1}(1-t)^{k-2} \left[t(1-t) - \frac{k-1}{2(2k-1)} \right] dt > 0.$$

Входящий в неравенство интеграл вычисляется. Будем искать его значение в виде

$$\int\limits_0^p t^{k-1}(1-t)^{k-2} \left[t(1-t) - \frac{k-1}{2(2k-1)} \right] dt = p^k q^{k-1} F(p),$$

где $F(p)$ - неизвестная функция. Дифференцируя по p обе части равенства, получим:

$$p^k q^{k-1} - p^{k-1} q^{k-2} \frac{k-1}{2(2k-1)} = kp^{k-1}q^{k-1}F(p) - (k-1)\,p^k q^{k-2} F(p) + p^k q^{k-1} F'(p).$$

После сокращения обеих частей на $p^{k-1}q^{k-2}$, придем к следующему дифференциальному уравнению первого порядка для неизвестной функции $F(p)$:

$$pq\left[F'(p) - 1\right] + \left[(2k-1)\,q - (k-1)\right] F(p) = -\frac{k-1}{2(2k-1)}.$$

Поищем решение уравнения в виде линейной функции $F(p) = ap + b$. Подстановка этого выражения дает:

$$(2ka-1)\,pq + (2k-1)\,qb - pa(k-1) - (k-1)\,b = -\frac{k-1}{2(2k-1)}.$$

Приравнивая слева и справа коэффициенты при pq (т.е. при p^2) и при p, найдем:

$$2ka - 1 = 0, \quad -(k-1)\,a - (2k-1)\,b = 0,$$

откуда

$$a = \frac{1}{2k}, \quad b = -\frac{k-1}{2k(2k-1)}.$$

Сравним оставшиеся свободные члены (не содержащие p) в левой и правой частях уравнения:

$$kb = -\frac{k-1}{2\left(2k-1\right)} \Longrightarrow b = -\frac{k-1}{2k\left(2k-1\right)}.$$

Это согласуется с выражением для b. Таким образом, решением дифференциального уравнения является функция

$$F\left(p\right) = \frac{1}{2k}\left(p - \frac{k-1}{2k-1}\right).$$

Подставляя найденное значение интеграла в неравенство, придем к

$$\frac{1}{2k}p^k q^{k-1}\left(p - \frac{k-1}{2k-1}\right) > 0$$

или

$$p > \frac{k-1}{2k-1}.$$

Получим отсюда ограничение на количество партий $n = 2k$:

$$2kp - p > k - 1 \Longrightarrow k\left(1-2p\right) < 1 - p \Longrightarrow 2k < \frac{2-2p}{1-2p}.$$

Таким образом, окончательно

$$n < 1 + \frac{1}{1-2p}.$$

Определим, при каких вероятностях p оптимальной является игра из двух партий. Для этого сравним правую часть неравенства с 4 :

$$\frac{1}{1-2p} < 3 \Longrightarrow p < \frac{1}{3}.$$

Итак, подобная стратегия оптимальна, если игрок A играет как минимум в два раза хуже игрока B. Легко показать, что для $1/3 \leq p < 2/5$ оптимальной уже является игра из четырех партий. Если же p близко к 0.5, т.е. игроки приблизительно равны по силам, то n может быть велико. Так, при $p = 0.45$ следует играть уже 10 партий.

Полиномиальный закон.

Обобщением схемы Бернулли является следующая схема испытаний. Производится n независимых испытаний и рассматривается полная группа попарно несовместных событий $\{H_1, H_2, \ldots, H_k\}$. Вероятность любого из событий $P\left(H_i\right) = p_i$ и не зависит от номера испытания. Определим вероятность того, что в n испытаниях событие H_1 произойдет m_1 раз, событие H_2 - m_2 раз, и т.д., событие H_k - m_k раз. Поскольку в каждом испытании осуществляется лишь какое-то одно из событий $H_1, H_2, \ldots, H_k$

$$\sum_{i=1}^{k} m_i = n.$$

Рассматриваемому событию благоприятствуют определенные последовательности произведений событий $H_1, H_2, \ldots, H_k$, простейшей из которых является

$$\underbrace{H_1 \ldots H_1}_{m_1} \cdot \underbrace{H_2 \ldots H_2}_{m_2} \cdot \ldots \cdot \underbrace{H_k \ldots H_k}_{m_k}.$$

В силу независимости испытаний вероятность появления такой последовательности равна

$$p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k}.$$

Поскольку все остальные благоприятствующие последовательности получаются из указанной всевозможными попарными перестановками событий H_i, вероятность их появления остается такой же. Значит, по теореме сложения вероятностей искомая вероятность получается умножением вероятности $p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k}$ на число благоприятствующих последовательностей событий H_i.

Подсчитаем это число на основе простейшей из последовательностей. Число всевозможных перестановок событий в этой последовательности равно $n!$, но поскольку, поменяв местами два одинаковых события H_i, мы не получаем новой последовательности, то число $n!$ необходимо уменьшить в $m_1! m_2! \ldots m_k!$ раз. Таким образом, окончательно для искомой вероятности получаем:

$$P_{m_1, m_2, \ldots, m_k; n} = \frac{n!}{m_1! m_2! \ldots m_k!} p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k}.$$

Эта формула называется *полиномиальным законом* или *вероятностью сложного события*.

Покажем, как отсюда получается биномиальный закон. Пусть требуется определить вероятность того, что событие A, вероятность появления которого в отдельном испытании равна p, произойдет k раз в n испытаниях. Рассмотрим полную группу попарно несовместных событий $A, \bar{A}$. Тогда, подставляя в полиномиальный закон $p_1 = P(A) = p, p_2 = P\left(\bar{A}\right) = 1 - p = q, m_1 = k, m_2 = n - k$, придем к

$$P_{k;n} = \frac{n!}{k!\,(n-k)!} p^k q^{n-k} = C_n^k p^k q^{n-k}.$$

15.Предельные теоремы в схеме Бернулли.

В приложениях часто приходится вычислять вероятности различных событий, связанных с числом успехов в n испытаниях Бернулли, при больших значениях n. В этом случае вычисления по формуле

$$P_{k;n} = C_n^k p^k q^{n-k}$$

становятся весьма затруднительными. Ситуация еще более усложняется, когда приходится суммировать указанные вероятности. Поэтому широкое распространение на практике получили приближенные асимптотические формулы.

Теорема Пуассона (для редких событий).

Если $n \to \infty$ и $p \to 0$ так, что $np \to \lambda \quad (0 < \lambda < \infty)$, то

$$P_{k;n} = C_n^k p^k q^{n-k} \to P_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$$

при любом постоянном $k : k = 0, 1, 2, \ldots$.

Доказательство: рассмотрим предел $P_{k;n}$ при $n \to \infty$.

$$\lim_{n\to\infty} P_{k;n} = \lim_{n\to\infty} C_n^k p^k q^{n-k} = \lim_{n\to\infty} \frac{n!}{k!\,(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n\,(n-1) \cdot \ldots \cdot (n-k+1)}{n \cdot n \cdot \ldots \cdot n} \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \ldots \left(1 - \frac{k-1}{n}\right) \frac{(1-\lambda/n)^n}{(1-\lambda/n)^k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^k}{k!} e^{-\lambda}. \qquad \circledast$$

Итак, в предельном случае относительно редких событий для вероятности k успехов получается следующая формула:

$$P_k = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Это - так называемый *закон Пуассона.* На практике данную формулу применяют для расчетов $P_{k;n}$ для больших, но конечных n и малых p.

Пример. Рассмотрим пример на схему Бернулли и закон Пуассона. Требуется узнать, сколько раз нужно сыграть в лотерею "Русское лото", покупая по одному билету в каждом тираже, чтобы вероятность Вашего выигрыша превысила заданное значение α ?

Оценим сначала величину p - вероятность выигрыша по одному билету в отдельном тираже. Если не учитывать всякие мелкие детали, то Ваш билет выиграет что-то (от 100 руб. до 350 000 руб.), если на нем закроются все 30 номеров. В обычном тираже не выпадают 8 бочонков из 90. Поэтому по классической схеме

$$p = \frac{C_{60}^{52}}{C_{90}^{82}} = \frac{60 \cdot 59 \cdot \ldots \cdot 54 \cdot 53}{90 \cdot 89 \cdot \ldots \cdot 84 \cdot 83} \simeq 0.033 = \frac{1}{30}.$$

Как видим, вероятность успеха мала, и можно пользоваться для подсчета как непосредственно формулой Бернулли, так и формулой Пуассона.

Формула Бернулли.

$$P_\alpha = 1 - q^n > \alpha \Longrightarrow q^n < 1 - \alpha \Longrightarrow n \ln q < \ln(1-\alpha) \Longrightarrow$$

$$\Longrightarrow n > \frac{\ln(1-\alpha)}{\ln q}.$$

Формула Пуассона.

$$P_\alpha = 1 - e^{-\lambda} > \alpha \Longrightarrow e^{-\lambda} < 1 - \alpha \Longrightarrow -\lambda < \ln(1-\alpha) \Longrightarrow$$

$$\Longrightarrow \lambda > -\ln(1-\alpha) \Longrightarrow n > -\frac{\ln(1-\alpha)}{p}.$$

Отличие в знаменателях выражений, но поскольку $p \ll 1$

$$\ln q = \ln(1-p) \simeq -p - \frac{p^2}{2} \simeq -p,$$

и обе формулы дают один и тот же результат. Поэтому для решения задачи можно применять любую схему. Так для $\alpha = 0.5$ $n > 21$, для $\alpha = 0.9$ $n > 69$.

Заметим, что при малых q можно пользоваться формулой Пуассона для числа неудач. Если же p и q заметно отличаются от нуля, то применяют локальную и интегральную теоремы Муавра-Лапласа.

Локальная теорема Муавра-Лапласа.

Если $n \to \infty$, а вероятность p постоянна $(0 < p < 1)$, величина $x_k = (k - np)/\sqrt{npq}$ равномерно ограничена по k и n $(-\infty < a \le x_k \le b < +\infty)$, то:

$$P_{k;n} \approx \frac{1}{\sqrt{2\pi npq}} \exp\left\{-\frac{(k-np)^2}{2npq}\right\}.$$

Доказательство: воспользуемся асимптотической формулой Стирлинга:

$$n! \simeq \left(\frac{n}{e}\right)^n \sqrt{2\pi n},$$

откуда следует, что

$$\ln n! \simeq n \ln n - n + \ln\sqrt{2\pi n}.$$

Прологарифмируем формулу для биномиального распределения.

$$\ln P_{k;n} = \ln C_n^k + k \ln p + (n-k)\ln q.$$

Рассмотрим отдельно

$$\ln C_n^k = \ln n! - \ln k! - \ln(n-k)!$$

Поскольку, $k = np + x_k\sqrt{npq}$, то $n - k = n - np + x_k\sqrt{npq} = nq - x_k\sqrt{npq}$. В силу конечности p и q и равномерной ограниченности x_k при $n \to \infty$: $k \to \infty$ и $(n-k) \to \infty$. В результате для вычисления $\ln C_n^k$ мы имеем право применить три раза вышеупомянутую формулу Стирлинга:

$$\ln C_n^k \simeq n\ln n - n + \ln\sqrt{2\pi n} - k\ln k + k - \ln\sqrt{2\pi k} -$$

$$-(n-k)\ln(n-k) + (n-k) - \ln\sqrt{2\pi(n-k)} =$$

$$= \ln\sqrt{\frac{n}{2\pi k(n-k)}} + n\ln n - k\ln k - (n-k)\ln(n-k).$$

Подставим все в исходное соотношение

$$\ln P_{k;n} \cong \ln \sqrt{\frac{n}{2\pi npnq}} + n\ln n - k\ln\frac{k}{p} - (n-k)\ln\frac{n-k}{q} =$$

$$= \ln\left(\frac{1}{\sqrt{2\pi npq}}\right) + n\ln n - k\ln\left[n\left(1 + x_k\sqrt{\frac{q}{np}}\right)\right] - (n-k)\ln\left[n\left(1 - x_k\sqrt{\frac{p}{nq}}\right)\right] =$$

$$= \ln\left(\frac{1}{\sqrt{2\pi npq}}\right) - np\left(1 + x_k\sqrt{\frac{q}{np}}\right)\ln\left(1 + x_k\sqrt{\frac{q}{np}}\right) -$$

$$-nq\left(1 - x_k\sqrt{\frac{p}{nq}}\right)\ln\left(1 - x_k\sqrt{\frac{p}{nq}}\right).$$

Воспользуемся приближенной формулой:

$$\ln(1+z) \cong z - \frac{z^2}{2} \text{ при } |z| \ll 1.$$

Тогда,

$$(1+z)\ln(1+z) \cong z(1+z)\left(1 - \frac{z}{2}\right) \cong z\left(1 + \frac{z}{2}\right).$$

Используя этот результат в формуле для $\ln P_{k;n}$, имеем:

$$\ln P_{k;n} \cong \ln\left(\frac{1}{\sqrt{2\pi npq}}\right) - npx_k\sqrt{\frac{q}{np}}\left(1 + \frac{x_k}{2}\sqrt{\frac{q}{np}}\right) + nqx_k\sqrt{\frac{p}{nq}}\left(1 - \frac{x_k}{2}\sqrt{\frac{p}{nq}}\right) =$$

$$= \ln\left(\frac{1}{\sqrt{2\pi npq}}\right) - x_k\sqrt{npq}\left(1 + \frac{x_k}{2}\sqrt{\frac{q}{np}} - 1 + \frac{x_k}{2}\sqrt{\frac{p}{nq}}\right) =$$

$$= \ln\left(\frac{1}{\sqrt{2\pi npq}}\right) - \frac{x_k^2}{2}(p+q) = \ln\left(\frac{1}{\sqrt{2\pi npq}}\right) - \frac{x_k^2}{2}.$$

Таким образом,

$$P_{k;n} \cong \frac{1}{\sqrt{2\pi npq}}\exp\left(-\frac{x_k^2}{2}\right). \qquad \boxtimes$$

Выведенную формулу часто используют для приближенных расчетов при $n > 100$ и при $npq > 20$. Ее применение оптимально при $p = q = 1/2$ (опыты с бросаниями симметричной монеты). При малых же p или q она дает большие ошибки.

В случае $p = q = 1/2$ и четного числа испытаний $2n$ имеем:

$$P_{k;2n} \cong \frac{1}{\sqrt{\pi n}}\exp\left\{-\frac{(k-n)^2}{n}\right\}.$$

Как видно, наиболее вероятным является выпадение одинакового числа "гербов" и "решек":

$$P_{n;2n} \cong \frac{1}{\sqrt{\pi n}}.$$

Интегральная теорема Муавра-Лапласа.

Если вероятность успеха p $(0 < p < 1)$ постоянна, то при $n \to \infty$

$$P\left\{a \leq \frac{\mu_n - np}{\sqrt{npq}} \leq b\right\} \to \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-x^2/2} dx$$

равномерно по $a, b \quad (-\infty \leq a \leq b \leq \infty)$, где μ_n - число успехов.

Доказательство: докажем теорему для фиксированных и конечных a и b. Очевидно,

$$P\left\{a \leq \frac{\mu_n - np}{\sqrt{npq}} \leq b\right\} = \sum_{a \leq x_k \leq b} P_{k;n},$$

где, как и в локальной теореме,

$$x_k = \frac{k - np}{\sqrt{npq}}.$$

Сумма

$$R_n = \sum_{x_k \in [a,b]} P_{k;n}$$

при $n \to \infty$ согласно локальной теореме Муавра-Лапласа приобретает вид

$$R_n \simeq \sum_{x_k \in [a,b]} \frac{1}{\sqrt{2\pi}} e^{-x_k^2/2} \cdot \frac{1}{\sqrt{npq}}.$$

Вычислим приращение

$$\Delta x_k = x_{k+1} - x_k = \frac{k + 1 - np}{\sqrt{npq}} - \frac{k - np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}.$$

В результате,

$$R_n \simeq \sum_{x_k \in [a,b]} \varphi(x_k)\, \Delta x_k,$$

где

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Таким образом, мы пришли к интегральной сумме Римана, которая при $n \to \infty$ $(\Delta x_k \to 0)$ сходится к интегралу

$$R_n \to \int\limits_a^b \varphi(x)\, dx = \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-x^2/2} dx. \qquad \boxtimes$$

Итак, при больших n справедлива приближенная формула

$$P\left\{a \leq \frac{\mu_n - np}{\sqrt{npq}} \leq b\right\} \simeq \Phi_0(b) - \Phi_0(a),$$

где

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2} dt.$$

Этот интеграл, называемый *интегралом ошибок*, не вычисляется в элементарных функциях, но для него существуют таблицы.

Историческая справка. Абрахам де Муавр (1667 – 1754) - одна из выдающихся фигур в теории вероятностей. Он родился во Франции, затем переехал в Англию. В 1718 г. была опубликована его основная работа "Доктрина шансов". Муавр скромно писал о своем открытии мирового значения: "Возьму на себя смелость утверждать, что это труднейшая проблема о случайном...". Результат он получил в 1733 г., а опубликован он был лишь в 1756 г. после его смерти.

Интегральная теорема Муавра-Лапласа позволяет оценить близость относительной частоты события к его вероятности. Если p - вероятность "успеха" в одном испытании, а μ_n - число "успехов" в n испытаниях, то необходимо оценить вероятность события

$$P\left\{\left|\frac{\mu_n}{n} - p\right| \le \Delta\right\}.$$

Тогда

$$P\left\{\left|\frac{\mu_n}{n} - p\right| \le \Delta\right\} = P\left\{\left|\frac{\mu_n - np}{n}\right| \le \Delta\right\} = P\left\{\left|\frac{\mu_n - np}{\sqrt{npq}}\right| \le \Delta\sqrt{\frac{n}{pq}}\right\} =$$

$$= P\left\{-\Delta\sqrt{\frac{n}{pq}} \le \frac{\mu_n - np}{\sqrt{npq}} \le \Delta\sqrt{\frac{n}{pq}}\right\} \simeq 2\Phi_0\left(\Delta\sqrt{\frac{n}{pq}}\right).$$

Здесь использована нечетность функции $\Phi_0(x)$.

Заметим, что поскольку $\Phi_0(0) = 0, \Phi_0(\infty) = 1/2$, в пределе $n \to \infty$ независимо от Δ

$$P\left\{\left|\frac{\mu_n}{n} - p\right| \le \Delta\right\} \to 1.$$

В этом смысле и следует понимать сходимость относительной частоты события $\nu_n(A)$ к его вероятности $P(A)$.