Содержание

В теории вероятности обычно изучают случайные события

Обычно наука занимается закономерностями, но так как в случайных экспериментах нет закономерностей, теория вероятности занимается поисков закономерности в сериях случайных экспериментах

Итак, в XVI веке начали с экспериментов бросков монеты:

число бросков	число гербов	частота
4040	2048	0.5069
12000	6019	0.5016
24000	12012	0.5005

Как можно видеть, частота стремится к 0.5 - появляется статистическая закономерность

1. Статистическое определение вероятности

Пусть проводится n реальных экспериментов, при которых событие A появилось n_A раз

Отношение $\dfrac{n_A}{n}$ называется частотой события A

Эксперименты показывают, что при увеличении числа n частота стабилизируется у некоторого числа, при котором мы понимаем статистическую вероятность: $P(A) \approx \dfrac{n_A}{n}$ при $n \to \infty$

Пространство элементарных исходов. Случайные события

Def. Пространством элементарных исходов Ω называется множество, содержащее все возможные исходы экспериментов, из которых при испытании происходит ровно один. Элементы этого множества называются элементарными исходами и обозначаются ω

Def. Случайными событиями называется подмножество $A \subset \Omega$. События A наступают, если произошел один из элементарных исходов из множества A

Ex. 1. Бросок монеты: $\Omega = \{\text{Г}, \text{Р}\}$, $A = \{\text{Г}\}$ - выпал герб

Ex. 2. Игральная кость: $\Omega = \{1, 2, 3, 4, 5, 6\}$, $A = \{\text{выпало четное число}\} = \{2, 4, 6\}$

Ex. 3. Монета бросается дважды.

а) Учитываем порядок: $\Omega = \{\text{ГГ}, \text{РР}, \text{РГ}, \text{ГР}\}$

а) Не учитываем порядок: $\Omega = \{\text{ГГ}, \text{РР}, \text{ГР}\}$

Ex. 4. Кубик дважды: $\Omega = \{\langle i, j\rangle \mid 1 \le i, j \le 6\}$

$A = \{\text{разность} \vdots 3\} = \{\langle 1, 4\rangle; \langle 4, 1\rangle; \langle 2, 5\rangle; \langle 5, 2\rangle; \dots\}$

Ex. 5. Монета бросается до первого герба: $\Omega = \{\text{Г}, \text{РГ}, \text{РРГ}, \dots\}$ - счетно-бесконечное множество

Ex. 6. Монета бросается на плоскость: $\Omega = \{\langle x, y\rangle \mid x, y \in \mathbb{R}, \langle x, y\rangle$ - центр монеты$\}$ - несчетное число исходов

Операции над событиями

Ω - достоверные события (наступают всегда)

$\varnothing$ - невозможное события (никогда не наступает, так как не содержит ни одного элем. исхода)

Введем операции:

Def. 1. Суммой $A+B$ называется событие, состоящее в том, что произошло события A или событие B (хотя бы одно из них)

Def. 2. Произведением $A \cdot B$ называется событие, состоящее в том, что произошло событие A и событие B (оба из них)

Nota. $A_1 + A_2 + \dots + A_n + \dots$ - произошло хотя бы одно из этих событий

$A_1 \cdot A_2 \cdot \dots \cdot A_n \cdot \dots$ - произошли все эти события

Def. 3. Противоположным A событием называется событие $\overline{A}$, состоящее в том, что событие A не произошло

Nota. $\overline{\overline{A}} = A$

Def. 4. Дополнение (разность) $A \setminus B$ называется событие $A \cdot \overline{B}$

Def. 5. События A и B называются несовместными, если их произведение - пустое множество (не могут произойти одновременно при одной эксперименте)

Def. 6. События A влечет события B, если $A \subset B$ (если наступает A, то наступит B)

Вероятность

Мы хотим присвоить какую-то числовую характеристику к каждому событию, отражающее его частоту наступления: $0 \le P(A) \le 1$ - вероятность наступления события A

Классическое определение вероятности

Пусть пространство случайных событий Ω содержит конечное число равновозможных исходов, тогда применимо классическое определение вероятности

Def. $\boxed{P(A) = \dfrac{|A|}{|\Omega|} = \dfrac{m}{n}}$, где n - число всех возможных исходов, m - число благоприятных исходов

В частности, если $\Omega = n$ и A_i - элем. исх., то $P(A_i) = \dfrac{1}{n}$

Свойства:

1) $0 \le P(A) \le 1$

2) $P(A) = 1 \quad (m = n)$

3) $P(\varnothing) = 0 \quad (m = 0)$

4) Если события A и B несовместны, то $P(A+B) = P(A) + P(B)$

Геометрическое определение вероятности (граф де Бюффон)

Пусть $\Omega \subset \mathbb{R}^n$ - замкнутая ограниченная область

$\mu(\Omega)$ - мера Ω в $\mathbb{R}^n$ (например, длина отрезка, площадь области на плоскости, объем тела в пространстве)

В эту область наугад бросаем точку. «Наугад» означает, что вероятность попадания в A зависит только от меры A и не зависит от ее расположения

В этом случае применимо геометрическое определение вероятности

$P(A) = \dfrac{\mu(A)}{\mu(\Omega)}$

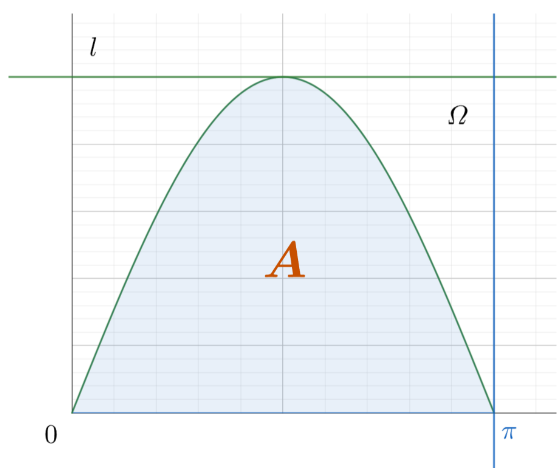
Ex. 1. Монета диаметром в 6 см бросается на пол, вымощенной квадратной плиткой со стороной 20 см, какова вероятность, что монета окажется целиком внутри одной плитки

$\mu(\Omega) = 20^2 = 400$

$\mu(A) = (20-3-3)^2 = 196$

$P(A) = \dfrac{\mu(A)}{\mu(\Omega)} = \dfrac{196}{400} = 0.49$

Ex. 2. Задача Бюффона об игле: пусть пол вымощен ламинатом, $2l$ - ширина доски, на пол бросается игла длины, равной ширине доски, найти вероятность того, что игла пересечет стык доски



Определим положение иглы координатами центра и углом, между иглой и стыком доски, причем можно считать, что эти величины независимы

$\exists x \in [0;1]$ - расстояние от центра до ближайшего края, $\varphi \in [0;\pi]$ - угол

$\Omega = [0;1] \times [0;\pi]$

Событие A (пересечет стык) наступает, если $x \le l \sin\varphi$

$P(A) = \dfrac{S(A)}{S(\Omega)}$

$S(\Omega) = \pi l$

$S(A) = \displaystyle\int_0^\pi l \sin\varphi d\varphi = -l\cos\varphi\Big|_0^\pi = -l(-1-1) = 2l$

$P(A) = \dfrac{2l}{\pi l} = \dfrac{2}{\pi}$

Построение модели случайных явлений

1. Задаем пространство элементарных исходов Ω

2. **Def.** Система $\mathcal{F}$ подмножеств Ω называется σ-алгеброй событий, если:
   1) $\Omega \in \mathcal{F}$;
   2) $A \in \mathcal{F} \Longrightarrow \overline{A} \in \mathcal{F}$;
   3) $A_1, A_2, \dots, A_n, \dots \in \mathcal{F} \Longrightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

   Свойства:
   (a) $\varnothing \in \mathcal{F}$, так как $\Omega \in \mathcal{F} \Longrightarrow \overline{\Omega} = \varnothing \in \mathcal{F}$
   (b) $A_1, A_2, \dots \in \mathcal{F} \Longrightarrow \bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

   $\square$ $A_1, A_2, \dots \in \mathcal{F} \Longrightarrow \overline{A}_1, \overline{A}_2, \dots \in \mathcal{F} \Longrightarrow \bigcup\limits_{i=1}^{\infty} \overline{A}_i \in \mathcal{F} \Longrightarrow \overline{\bigcup\limits_{i=1}^{\infty} \overline{A}_i} = \bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F}$ $\square$

   (c) $A, B \in \mathcal{F} \Longrightarrow A \setminus B \in \mathcal{F}$

   $\square$ $A, B \in \mathcal{F} \Longrightarrow A, \overline{B} \in \mathcal{F} \Longrightarrow A \setminus B = A \cdot \overline{B} \in \mathcal{F}$ $\square$

   Ex. 1. $\mathcal{F} = \{\varnothing, \Omega\}$

   Ex. 2. $\mathcal{F} = \{\varnothing, \Omega, A, \overline{A}\}$

   Ex. 3. **Def.** Борелевская σ-алгебра $\mathcal{B}(\mathbb{R})$ - минимальная σ-алгебра, содержащая все возможные интервалы на прямой

3. **Def.** $\sqsupset$ Ω - пространство элементарных исходов, $\mathcal{F}$ - его σ-алгебра событий. *Вероятностью* на $(\Omega, \mathcal{F})$ называется функция $P: \mathcal{F} \to \mathbb{R}$ со свойствами:
   (a) $P(A) \geq 0 \quad \forall A \in \mathcal{F}$ (неотрицательность)
   (b) Если $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$ - несовместное, то $P(\sum\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$ (свойство счетной аддитивности)
   (c) $P(\Omega) = 1$ (условие нормированности)

Def. Из этого тройка $(\Omega, \mathcal{F}, P)$ называется вероятностным пространством

Свойства вероятности

1. Так как $\varnothing$ и Ω - несовместные, то $1 = P(\Omega) = P(\Omega + \varnothing) = 1 + P(\varnothing) \Longrightarrow P(\varnothing) = 0$
2. Формула обратной вероятности: $P(A) = 1 - P(\overline{A})$

   $\square$ A и $\overline{A}$ - несовместные и $A + \overline{A} = \Omega \Longrightarrow P(A + \overline{A}) = P(\Omega) = 1$ $\square$

3. $P(A) = 1 - P(\overline{A}) \leq 1$

Аксиома непрерывности

Пусть имеется убывающая цепочка событий $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$ и $\bigcap\limits_{i=1}^{\infty} A_n = \varnothing$

Тогда $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$

При непрерывном изменении области $A \subset \Omega \subset \mathbb{R}^n$ соответствующая вероятность $P(A)$ также должна изменятся непрерывно

Th. Аксиома непрерывности следует из аксиомы счетной аддитивности

$\square$

Ясно, что $A_n = \sum\limits_{i=n}^{\infty} A_i\overline{A}_{i+1} + \prod\limits_{i=n}^{\infty} A_i$

$\prod\limits_{i=n}^{\infty} A_i = A_n \cdot \prod\limits_{i=n+1}^{\infty} A_i = \prod\limits_{i=1}^{n} \cdot \prod\limits_{i=n+1}^{\infty} A_i = \prod\limits_{i=1}^{\infty} = \varnothing \implies A_n = \sum\limits_{i=n}^{\infty} A_n\overline{A_{n+1}}$ и так как эти события несовместны, то по свойству счетной аддитивности $P(A_n) = \sum\limits_{i=n}^{\infty} P(A_i\overline{A_{i+1}})$ - это остаток (хвост) сходящегося ряда

$P(A_1) = \sum\limits_{i=1}^{\infty} P(A_i\overline{A_{i+1}}) = \sum\limits_{i=1}^{n-1} P(A_i\overline{A_{i+1}}) + P(A_n)$ и $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$ по необходимому признаку сходимости

$\square$

Nota. Аксиому счетной аддитивности можно вывести из конечной аддитивности и аксиомы счетной непрерывности

Свойства операций сложения и умножения

1. Свойство дистрибутивности: $A \cdot (B + C) = AB + AC$
2. Формула сложения: если A и B несовместны, то $P(A + B) = P(A) + P(B)$
3. Формула сложения вероятностей: $P(A + B) = P(A) + P(B) - P(AB)$

$\square$

$A + B = A\overline{B} + AB + \overline{A}B$ - несовместные события $\implies P(A + B) = P(A\overline{B}) + P(AB) + P(\overline{A}B) = (P(A\overline{B}) + P(\overline{A}B)) - P(AB) = P(A) + P(B) - P(AB)$

$\square$

Ex. Из колоды в 36 карт достали одну карту. Какова вероятность того, что будет дама или пика

Пусть Д - дама, П - пика, $P(\text{Д} + \text{П}) = P(\text{Д}) + P(\text{П}) - P(\text{ДП}) = \frac{4}{36} + \frac{9}{36} - \frac{1}{36} = \frac{1}{3}$

Формула сложения при $N = 3$: $P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_2A_3) - P(A_1A_3) - P(A_1A_2) + P(A_1A_2A_3)$

Общий случай: $P(A_1 + A_2 + \dots + A_n) = \sum\limits_{i=1}^{n} P(A_i) - \sum\limits_{i<j} P(A_iA_j) + \sum\limits_{i<j<k} P(A_iA_jA_k) + (-1)^{n-1} \cdot P(A_1A_2\dots A_n)$ - формула включения и исключения

Ex. n писем случайно раскладывается по n конвертам. Найти вероятность того, что хотя бы одно письмо окажется в своем конверте
$\sqsupset A_i$ - i-ое письмо в своем конверте
$P(A_i) = \dfrac{1}{n}; P(A_iA_j) = \dfrac{1}{A_n^2}; P(A_iA_jA_k) = \dfrac{1}{A_n^3}; P(A_1A_2\dots A_n) = \dfrac{1}{n!}$
Слагаемых вида A_i - n штук; A_iA_j - C_n^2; $A_iA_jA_k$ - C_n^3; $A_1A_2\dots A_n$ - 1 штука
$P(A) = P(A_1 + A_2 + \dots + A_n) = n \cdot \dfrac{1}{n} - C_n^2\dfrac{1}{A_n^2} + C_n^3\dfrac{1}{A_n^3} - \dots + (-1)^{n-1}\dfrac{1}{n!} = 1 - \dfrac{1}{2} + \dfrac{1}{3!} - \dots + (-1)^{n-1}\dfrac{1}{n!}$
Так как $e^{-1} = 1 - 1 + \dfrac{1}{2} - \dfrac{1}{3!} + \dots$, то при $n \to \infty \quad P(A) \underset{n\to\infty}{\to} 1 - e^{-1} \approx 0.63$
Независимые события
Под независимыми событиями логично подразумевать события, не связанные причинно-следственной связью (то есть когда факт наступления одного не влияет на оценку вероятности другого)
$\sqsupset |\Omega| = n; |A| = m_1; |B| = m_2$
Проведем пару независимых испытаний. Тогда получаем пространство элементарных исходов $\Omega \times \Omega$ и $|\Omega \times \Omega| = n^2$
По основному принципу комбинаторики $|A \cdot B| = m_1 \cdot m_2$
$P(AB) = \dfrac{|A \cdot B|}{|\Omega \times \Omega|} = \dfrac{m_1m_2}{n^2} = P(A) \cdot P(B)$

Def. События A и B называются независимыми, если $P(A \cdot B) = P(A) \cdot P(B)$
Lab. $\sqsupset P(A), P(B) \neq 0$, доказать, что если A и B несовместны, то они зависимы
Свойство: Если A и B независимы, то независимы $\overline{A}$ и $\overline{B}$, A и $\overline{B}$, $\overline{A}$ и B
Доказательство: $A = A \cdot (B + \overline{B}) = AB + A\overline{B}$ - несовместные события $\Longrightarrow P(A) = P(AB) + P(A\overline{B}) \Longrightarrow P(A\overline{B}) = P(A) - P(AB) = P(A) - P(A) \cdot P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B}) \Longrightarrow$ независимы

Def. События $A_1, A_2, \dots A_n$ - независимы в совокупности, если для любого набора $i_1, i_2, \dots, i_k$ $(2 \le k \le n)$ $P(A_{i_1} \cdot A_{i_2} \cdot \dots \cdot A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$

Nota. Из независимости в совокупности при $k = 2$ получаем попарную независимость. Обратное утверждение неверно

Ex. (С. Бернштейн)
Пусть имеется правильный тетраэдр, одна грань окрашена в красный, вторая в синий, третья в зеленый, а четвертая во все эти три цвета.
Подбросили тетраэдр, $\sqsupset A$ - грань, которая содержит красный цвет, B - синий, C - зеленый.
$P(A) = P(B) = P(C) = \dfrac{2}{4} = \dfrac{1}{2}$

Так как $P(AB) = P(AC) = P(BC) = \dfrac{1}{4}$

$P(AB) = \dfrac{1}{4} = \dfrac{1}{2} \cdot \dfrac{1}{2} = P(A)P(B)$ - попарная независимость

$P(ABC) = \dfrac{1}{4} \neq P(A)P(B)P(C)$ - но вот независимость в совокупности не соблюдается

Ex. (Шевалье де Мере, Паскаль, Ферма, $\approx$ 1650 г.)

Какова вероятность того, что при 4 бросании кости выпадет одна шестерка

A_1 - при первом броске шестерка, A_2 - при втором, A_3 - при третьем, A_4 - при четвертом

B - выпала хотя бы одна шестерка при 4 бросках

$B = A_1 + A_2 + A_3 + A_4$ - совместные события, но независимые

Найдем обратную вероятность: $\overline{B}$ - ни разу не выпала шестерка

$\overline{B} = \overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} \cdot \overline{A_4}$

$P(\overline{A_1}) = P(\overline{A_2}) = P(\overline{A_3}) = P(\overline{A_4}) = \dfrac{5}{6}$

$\overline{B} = P(\overline{A_1})P(\overline{A_2})P(\overline{A_3})P(\overline{A_4}) = \left(\dfrac{5}{6}\right)^4 \approx 0.482$

$P(B) = 1 - P(\overline{B}) \approx 0.52$

Условная вероятность

Условная вероятность $P(A|B)$ (или $P_B(A)$) - вероятность события A, вычисленная в предположении, что событие B уже произошло

Ex. Бросается кость один раз, известно, что выпало больше 3 очков. Найти вероятность того, что выпало четное число очков

A - выпало четное число очков

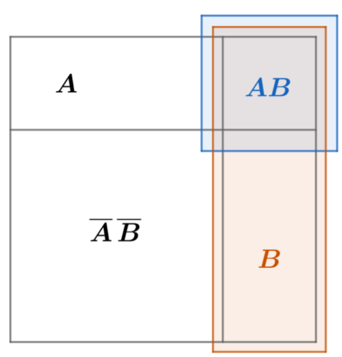
B - выпало больше трех очков

$\Omega = \{1, 2, 3, 4, 5, 6\}; |\Omega| = 6; A = \{2, 4, 6\}; B = \{4, 5, 6\}$

$P(A|B) = \dfrac{2}{3} = \dfrac{\frac{2}{6}}{\frac{3}{6}} = \dfrac{P(AB)}{P(B)}$

Интерпретация с помощью геометрической вероятности:

$P(A|B) = \dfrac{S_{AB}}{S_B} = \dfrac{\frac{S_{AB}}{S_\Omega}}{\frac{S_B}{S_\Omega}}$



Def. Условной вероятностью события A при условии, что имело место событие B, называется величина $P(A|B) = \dfrac{P(AB)}{P(B)}$

Ex. Известно, что среди населения 1% воров. В комнате, где находилось 10 гостей, у хозяина пропал кошелек. Какова вероятность того, что произвольный гость является вором.

A - гость является вором $P(A) = 0.01$

B - пропал кошелек (хотя бы один вор среди гостей есть)
$P(A|B) = \dfrac{P(AB)}{P(B)} = \dfrac{P(AB)}{1-P(\overline{B})} = \dfrac{P(A)}{1-0.99^{10}} = \dfrac{0.01}{1-0.99^{10}} = 0.105$
Формула умножения:

В качестве следствия условной вероятности получаем:
$P(A|B) = \dfrac{P(AB)}{P(B)} \Longrightarrow P(AB) = P(B)\cdot P(A|B) = P(A)\cdot P(B|A)$
Общий случай:

$P(A_1A_2A_3\dots A_n) = P(A_1)P(A_2|A_1)P(P_3|A_1A_2)\dots P(A_n|A_1A_2\dots A_{n-1})$

□

База индукции $P(AB) = P(B)P(A|B)$

Шаг индукции: пусть верно при $n-1$:

$P(A_1A_2A_3\dots A_{n-1}) = P(A_1)P(A_2|A_1)P(P_3|A_1A_2)\dots P(A_n|A_1A_2\dots A_{n-2})$

$P(A_1A_2A_3\dots A_n) = P(A_1A_2A_3\dots A_{n-1})\cdot P(A_n|A_1A_2\dots A_{n-1}) =$

$P(A_1)P(A_2|A_1)P(P_3|A_1A_2)\dots P(A_n|A_1A_2\dots A_{n-1})$

□

Ex. Студент выучил 1 билет из n, в группе n студентов. Каким по очереди ему нужно зайти, чтобы вероятность сдать экзамен была наибольшей

Пусть A_i - билет, вытянутый на i-ом шаге $(1 \le i \le n)$

A - студент сдал экзамен

$P(A) = P(\overline{A_1}\cdot\overline{A_2}\cdot\dots\cdot\overline{A_{i-1}}\cdot A_i) = \dfrac{n-1}{n}\cdot\dfrac{n-2}{n-1}\cdot\dots\cdot\dfrac{n-(i-1)}{n-(i-2)}\cdot\dfrac{1}{n-(i-1)} = \dfrac{1}{n}$

Полная группа событий

Def. События $H_1, H_2, \dots, H_n, \dots$ образуют полную группу событий, если они попарно несовместны и содержат все возможные элементарные исходы

$H_i \cap H_j = \varnothing\ \forall i, j \qquad \bigcup\limits_{i=1}^{\infty} H_i = \Omega$

Следствие: $\sum\limits_{i=1}^{\infty} P(H_i) = 1$

Th. Формула полной вероятности. $\exists H_1, H_2, \dots, H_n, \dots$ - полная группа событий. Тогда $P(A) = \sum\limits_{i=1}^{\infty} P(H_i)P(A|H_i)$

□

$P(A) = P(\Omega A) = P((H_1 + H_2 + H_3 + \dots)A) = P(H_1A + H_2A + H_3A + \dots) = [H_i \cdot A \cdot H_j \cdot A = \varnothing \cdot A] = P(H_1A) + P(H_2A) + \dots = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots$

□

Th. Формула Байеса. $\sqsupset H_1, H_2, \dots, H_n$ - полная группа событий, и известно, что событие A уже произошло

Тогда $P(H_k|A) = \dfrac{P(H_k)P(A|H_k)}{\sum_{i=1}^{\infty} P(H_i)P(A|H_i)}$

□

$P(H_k|A) = \dfrac{P(H_kA)}{P(A)} = \dfrac{P(H_k)P(A|H_k)}{\sum_{i=1}^{\infty} P(H_i)P(A|H_i)}$

□

Ex. 1. В первой коробке 4 белых и 2 черных шара, во второй 1 белый и 2 черных. Из первой коробки во вторую переложили 2 шара, затем из второй коробки достали шар. Какова вероятность того, что он оказался белым

$\sqsupset H_1$ - переложили 2 белых H_2 - 2 черных

H_3 - разного цвета

A - из второй коробки достали белый шар

$P(H_1) = \frac{4}{6} \cdot \frac{3}{5} = \frac{6}{15}$

$P(H_2) = \frac{2}{6} \cdot \frac{1}{5} = \frac{1}{15}$

$P(H_3) = \frac{4}{6} \cdot \frac{2}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{4}{15} + \frac{4}{15} = \frac{8}{15}$

$P(A) = P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + P(H_3) \cdot P(A|H_3) = \frac{6}{15} \cdot \frac{3}{5} + \frac{1}{15} \cdot \frac{1}{5} + \frac{8}{15} \cdot \frac{2}{5} = \frac{18}{75} + \frac{1}{75} + \frac{16}{75} = \frac{35}{75} = \frac{7}{15}$

Ex. 2. Вероятность попадания первого стрелка в цель 0.9, а второго 0.3. Наугад вызванный стрелок попал в цель. Какова вероятность того, что это бы первый стрелок?

H_1 - вызван первый стрелок

H_2 - вызван второй стрелок

A - стрелок попал

$P(H_1) = P(H_2) = \frac{1}{2}$

$P(A|H_1) = 0.9 \qquad P(A|H_2) = 0.3$

$P(H_1|A) = \dfrac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)|P(A|H_2)} = \dfrac{\frac{1}{2}0.9}{\frac{1}{2}0.9 + \frac{1}{2}0.3} = \dfrac{9}{9+3} = 0.75$

Ex. 3. По статистике раком болеет 1% населения. Тест дает правильный результат в 99% случаев. Тест оказался положительный. Найти вероятность того, что человек болен.

H_1 - человек болен

H_2 - человек здоров

A - анализ положительный

$P(H_1) = 0.01$

$P(H_2) = 0.99$

$P(A|H_1) = 0.99$

$P(A|H_2) = 0.01$

$P(H_1|A) = \dfrac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \dfrac{0.01 + 0.99}{0.01 \cdot 0.99 + 0.99 \cdot 0.01} = \dfrac{1}{2} = 0.5$

Допустим, что второй независимый с первым анализ также оказался положительным. Найти вероятность того, что человек болен.

$P(H_1) = 0.01 \qquad P(H_2) = 0.99$

$P(AA|H_1) = 0.99^2 \qquad P(AA|H_2) = 0.01^2$

$P(H_1|AA) = \dfrac{0.01 + 0.99^2}{0.01 \cdot 0.99^2 + 0.99 \cdot 0.01^2} = \dfrac{0.99}{0.99 + 0.01} = 0.99$

Интуитивно вероятность $\dfrac{1}{2}$ может поддаваться непониманию, однако можно рассуждать так: пусть в городе живут 10000 человек, из них 100 болеют, а у 99 из них положительный анализ; у других 9900 положительный анализ всего лишь у 99, отсюда выходит $\dfrac{1}{2}$

Ex. 4. В телевизионной студии 3 двери 🚪🚪🚪, за одной из них приз 🚗. Игрок выбрал наугад одну из 3 дверей, после чего ведущий открывает одну из двух оставшихся дверей и показывает, что там приза нет 🛴. После чего предлагает игроку поменять свой выбор. Стоит ли игроку соглашаться?

H_1 - игрок угадал

H_2 - игрок не угадал

A - ведущий открыл дверь без приза

$P(H_1) = \dfrac{1}{3} \qquad P(H_2) = \dfrac{2}{3}$

$P(A|H_1) = 1 \qquad P(A|H_2) = \dfrac{1}{2}$

$P(H_1|A) = \dfrac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot \frac{1}{2}} = \dfrac{1}{2}$

Но это неправильно, так как действия ведущего неслучайны - он всегда откроет дверь без приза

В этом случае, если мы гипотетически выберем 300 дверей, в 100 случаях мы отгадаем, ведущий откроет любую дверь без приза; но в 200 случаях мы не отгадаем, ведущий откроет вторую дверь без приза, и в этом случае мы сможем поменяться на дверь с призом, отсюда шанс $\dfrac{2}{3}$, если мы поменяем свой выбор

Ex. 5. Вероятность того, что в семье с детьми ровно k детей, равна $\frac{1}{2^k}$, $k = 1, 2, \ldots$ Какова вероятность того, что в семье один мальчик, если известно, что нет девочки? Рождения мальчиков и девочек равновероятны.
H_i - в семье i детей $(1 \le i < \infty)$
$P(H_i) = \frac{1}{2^i}$
A - в семье нет девочки
$P(A|H_1) = \frac{1}{2}$
$P(A|H_2) = \frac{1}{4}$
$P(A|H_i) = \frac{1}{2^i}$
$P(H_1|A) = \frac{\frac{1}{2}\frac{1}{2}}{\sum_{i=1}^{\infty} \frac{1}{2^i} \cdot \frac{1}{2^i}} = \frac{\frac{1}{4}}{\frac{\frac{1}{4}}{1-\frac{1}{4}}} = \frac{3}{4} = 0.75$

Серия испытаний Бернулли

Схемой Бернулли - называется серия одинаковых независимых экспериментов, каждый из которых имеет 2 исхода: произошло интересующее нас событие или нет
$p = p(A)$ - вероятность успеха при одном испытании
$q = 1 - p$ - вероятность неудачи
v_n - число успехов в серии из n испытаний
$p(v_n = k) = p_n(k)$
Из этого получаем формулу Бернулли:

Th. Вероятность того, что при n испытаниях произойдет ровно k успехов, равна
$p_n(k) = C_n^k p^k q^{n-k}$

□
Рассмотрим один из элементарных исходов, благоприятных данному событию:
$A_n = \underbrace{\text{УУУ} \ldots \text{У}}_{k}\underbrace{\text{Н} \ldots \text{ННН}}_{n-k}$ - k успехов, $n - k$ неудачи
$p(\text{У}) = p, p(\text{Н}) = q$
Так как испытания независимы, то $p(A_n) = p^k q^{n-k}$
Остальные элементарные исходы имеют ту же вероятность, перебираем все расстановки исходов, получаем C_n^k, в итоге, получаем формулу Бернулли
□

Ex. Вероятность попадания стрелка при одном выстреле - 0.8. Какова вероятность того, что из пяти выстрелов точными будут три

$n = 5 \quad p = 0.8 \quad q = 1 - p = 0.2 \quad k = 3$

$p_5(3) = C_5^3 p^3 q^2 = 0.2048$

Наиболее вероятное число успехов

Выясним, при каком значении k вероятность предшествующего числа успехов $k-1$ будет не более, чем вероятность k успехов

$p_n(k-1) \le p_n(k)$

$C_n^{k-1} p^{k-1} q^{n-k+1} \le C_n^k p^k q^{n-k}$

$\frac{n!}{(k-1)!(n-k+1)!} q \le \frac{n!}{(k)!(n-k)!} p$

$\frac{q}{(k-1)!(n-k+1)!} \le \frac{p}{(k)!(n-k)!}$

$\frac{q}{n-k+1} \le \frac{p}{k}$

$k(1-p) \le p(n-k+1)$

$k \le np + p$

Отсюда $np + p - 1 \le k \le np + p$

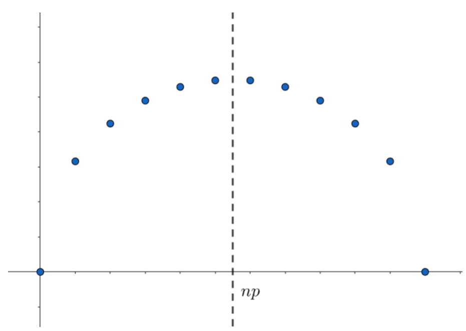
Рассмотрим 3 ситуации:

1) np - целое, тогда $np + p$ - нецелое, и $k = np$ - наиболее вероятное

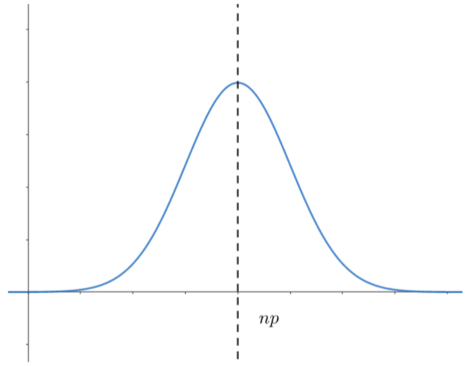
2) $np + p$ - нецелое, тогда $k = \lfloor np + p \rfloor$

3) $np + p$ - целое, тогда $np + p - 1$ - целое, и 2 наиболее вероятных числа успеха

Геометрическая интерпретация:



При увеличении числа n точки превращаются в кривую Гаусса



При увеличении числа испытаний n формула Бернулли вырождается в следующие асимптотические формы (применяем, если требуется найти вероятность точного числа успеха)

1) локальная формула Муавра-Лапласа

$p_n(k) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{npq}}\varphi(x)$, где $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ - функция гаусса

$x = \frac{k-np}{\sqrt{npq}}$

Свойства $\varphi(x)$:

1. $\varphi(x) = \varphi(-x)$ - функция четная

2. при $x > 5 \quad \varphi(x) \approx 0$

2) Интегральная формула Муавра-Лапласа (если требуется найти вероятность того, что число успехов в данном диапазоне)

$p_n(k_1 \le k \le k_2) \xrightarrow[n\to\infty]{} \Phi(x_2) - \Phi(x_1)$, где $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$ - функция Лапласа

$x_1 = \frac{k_1-np}{\sqrt{npq}}$ - отклонение от левой границы, $x_2 = \frac{k_2-np}{\sqrt{npq}}$ - отклонение от правой

Свойства $\Phi(x)$

1. $\Phi(-x) = -\Phi(x)$ - функция нечетная

2. при $x > 5 \quad \Phi(x) \approx 0.5$

Nota. Эти формулы обычно можно применять при $n \ge 100$ и $0.1 \le p \le 0.9$

Nota. В некоторых источниках под функцией Лапласа подразумевают другую функцию: $F_0(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{t^2}{2}}dt$ - стандартное отклонение. Эта функция отличается от $F_0 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^0 e^{-\frac{t^2}{2}}dt + \Phi(x) = \frac{1}{2} + \Phi(x)$

Так как $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ - интеграл Пуасона

Ex. Вероятность попадания стрелка в цель 0.8, стрелок сделал 400 выстрелов. Найти вероятность того, что:

а) произошло ровно 330 попаданий

б) произошло от 312 до 336 попаданий

а) $x = \frac{k-np}{\sqrt{npq}} = \frac{330-400\cdot 0.8}{\sqrt{400\cdot 0.8\cdot 0.2}} = \frac{330-320}{8} = 1.25$

$p_{400}(330) \approx \frac{1}{\sqrt{npq}}\varphi(1.25) = \frac{1}{8}\varphi(1.25) \approx \frac{1}{8}\cdot 0.1826 \approx 0.0228$

б) $x_1 = \frac{312-320}{8} = -1$, $x_2 = \frac{336-320}{8} = 2$

$p_{400}(312 \le k \le 336) \approx \Phi(2) - \Phi(-1) = \Phi(2) + \Phi(1) \approx 0.4772 + 0.3413 = 0.8185$

Статистическое понятие вероятности

Пусть проводим n реальных экспериментов, n_A - число появления события A, $\frac{n_A}{n}$ - относительная частота события A.
Эксперименты с монетой показали, что при больших n, $\frac{n_A}{n} \approx p(A)$ - явление стабилизации
Вероятность отклонения относительной частоты от вероятности события
n - число испытаний, $p = p(A)$, $\frac{n_A}{n}$ - экспериментальная частота
$p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) = p\left(-\varepsilon \le \frac{n_A}{n} - p \le \varepsilon\right) = p(-n\varepsilon \le n_A - np \le n\varepsilon) = p(np - n\varepsilon \le n_A \le n\varepsilon + np) \underset{n\to\infty}{\longrightarrow}$ [по интегральной формуле Лапласа] $\underset{n\to\infty}{\longrightarrow} \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) - \Phi\left(-\frac{n\varepsilon}{\sqrt{npq}}\right)$
$= \Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) - \Phi\left(-\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$
$= 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$
Итак, получили, что нужная нам вероятность $p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \approx 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$

Закон больших чисел Бернулли

Итак, $p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 2\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right)$
при $n \to \infty, \sqrt{n} \to \infty, \frac{\varepsilon}{\sqrt{pq}}\sqrt{n} \to \infty, \Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) \to 0.5, p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \to 2 \cdot 0.5 = 1$ - закон больших чисел показывает, что вероятность попадания относительной частоты в ε-трубу вероятность события приближается к 1
$\lim\limits_{n\to\infty} p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) = 1$ или $\frac{n_A}{n} \underset{n\to\infty}{\longrightarrow} p$ - сходимость по вероятности

Ex. Для оценки доли p курящих людей берется выборка объема n, и делается оценка доли курящих людей по формуле $p^* = \frac{n_A}{n}$. Каким должен быть объем n, чтобы с вероятностью $\gamma = 0.95$ данная оценка отличалась от истинного значения не более, чем на $\varepsilon = 0.01$
По формуле вероятности отклонения частоты от вероятности $p(|p^* - p| \le \varepsilon) = p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \approx$
$2\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) = 0.95$
$\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) = 0.475$
$\frac{\varepsilon}{\sqrt{pq}}\sqrt{n} = 1.96$
$\frac{1}{\sqrt{pq}}\sqrt{n} = 196$
$\frac{n}{pq} = 38416$
$n \ge 38416pq$

В самое худшей ситуации $pq \le 0.5^2 = \dfrac{1}{4}$
$n \ge \dfrac{38416}{4} = 9604$

Схема испытаний и соответствующее распределение

Введем обозначения:
n - число испытаний
p - вероятность успеха при одном испытании
$q = 1 - p$ - вероятность неудачи

I. Схема Бернулли

$\exists v_n$ - число успехов в серии из n испытаний
$P_n(v_n = k) = C_n^k p^k q^{n-k}, \qquad k = 0, 1, \ldots, n$

Def. Соответствие $k \to C_n^k p^k q^{n-k}, \quad k = 0, \ldots, n$ называется биномиальным распределением (обозначается $B_{n,p}$ или $B(n, p)$)

II. Схема до первого успешного испытания

Пусть проводится бесконечная серия испытаний, которая заканчивается после первого успешного испытания под номером τ

Th. $P(\tau = k) = q^{k-1}p, \qquad k = 1, 2, \ldots$

$\square$

$P(\tau = k) = P(\underbrace{\text{Н}\ldots\text{Н}}_{k-1}\text{У}) = q^{k-1}p$

$\square$

Def. Соответствие $k \to q^{k-1}p, k \in \mathbb{N}$ называется геометрическим распределение вероятности (обозначается G_p или $G(p)$)

Nota. Геометрическое распределение обладает свойством нестарения или свойством отсутствия последействия

Th. $\exists P(\tau = k) = q^{k-1}p, k \in \mathbb{N}$. Тогда $\forall n, k \ge 0 \quad P(\tau > n + k \mid \tau > n) = P(\tau > k)$

□

Заметим, что $P(\tau > m) = q^m$, первые m - неудачи

$P(\tau > n+k|\tau > n) = \dfrac{P(\tau > n+k, \tau > n)}{P(\tau > n)} = \dfrac{P(\tau > n+k)}{P(\tau > n)} = \dfrac{q^{n+k}}{q^n} = q^k$

□

Nota. $P(\tau = n+k \mid \tau > n) = p(\tau = k)$ - Lab. доказать

III. Схема испытаний с несколькими исходами

Пусть при n независимых испытаний могут произойти m исходов (несовместных)

p_i - вероятность i-ого исхода при одном испытании

Th. Вероятность того, что при n испытаниях первый исход появится n_1 раз, второй - n_2 раз, m-ый - n_m $(\sum\limits_{i=1}^{m} n_i = n)$ равно

$P_n(n_1, n_2, \ldots, n_m) = \dfrac{n!}{n_1!n_2!\ldots n_m!} p_1^{n_1} p_2^{n_2} \ldots p_m^{n_m}$

При $m = 2$ получаем формулу Бернулли

□

Рассмотрим следующий благоприятный исход, обозначим A_1

$A_1 = \underbrace{11\ldots1}_{n_1}\underbrace{22\ldots2}_{n_2}\ldots\underbrace{mm\ldots m}_{n_m}$

$p(A_1) = p_1^{n_1} p_2^{n_2} \ldots p_m^{n_m}$

Все остальные благоприятные исходы имеют ту же вероятность и отличаются лишь расположением i-ых исходов на n позициях, получаем мультиномиальную теорему: $\dfrac{n!}{n_1!n_2!\ldots n_m!}$

В итоге получаем требуемую формулу

□

Ex. Два одинаковых сильных шахматиста играют шесть партий

Вероятность ничьи в партии - 0.5. Какова вероятность того, что второй игрок выиграет две партии, а еще три сведет к ничьей

1-ый исход - выиграл 1 игрок

2-ой исход - выиграл 2 игрок

3-ий исход - ничья

$n = 6; \quad p_3 = 0.5; \quad p_1 = p_2 = \dfrac{1 - p_3}{2} = 0.25$

$P_6(1;2;3) = \frac{6!}{1!2!3!}\left(\frac{1}{4}\right)^1\left(\frac{1}{4}\right)^2\left(\frac{1}{2}\right)^3 = \frac{4\cdot 5\cdot 6}{2}\frac{1}{2^9} \approx 0.12$

IV. Урновая схема

В урне N шаров, из которых K шаров белые, $N-K$ - черные

Из урны вынимаем (без учета порядка) n шаров. Найти вероятность, что из них k белых

а) Схема с возвратом (после каждого раза кладем шар обратно). В этом случае вероятность вынуть белый шар одинакова и равна $\frac{K}{N}$. Получаем схему Бернулли: $P_n(k) = C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k}$

б) Схема без возврата - вынутый шар мы выбрасываем

$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}$

Def. Соответствие $k \to \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}, k = 0, \dots, n$ называется гипергеометрическим распределением

Nota. Если $K, N \to \infty$ так, что $\frac{K}{N} \approx p$ (не меняется), а n и k зафиксировать, то после выбора n шаров пропорции состава шаров не сильно изменятся, поэтому логично предположить, что гипергеометрическое распределение будет сходиться к биномиальному

Th. Если $K, N \to \infty$ таким образом, что $\frac{K}{N} \to p \in (0;1)$, а n и $0 \le k \le n$ фиксированы, то вероятность при гипергеометрическом распределении будет стремиться к биномиальному:

$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \to C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k}$

Воспользуемся леммой: $C_n^k \sim \frac{n^k}{k!}$ при $n \to \infty$ и фиксированном k

Доказательство леммы: $C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{n^k}\frac{n^k}{k!} = 1\left(1-\frac{1}{n}\right)\dots\left(1-\frac{k-1}{n}\right)\frac{n^k}{k!} \sim \frac{n^k}{k!}$

□

$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \sim \frac{K^k}{k!}\frac{(N-K)^{n-k}}{N^n}\frac{n!}{N^n} = \frac{n!}{k!}\frac{(N-K)^{n-k}}{N^n}\frac{K^k}{N^n} = C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k} \to C_n^k\left(\frac{K}{N}\right)^k\left(1-\frac{K}{N}\right)^{n-k}$

□

V. Схема Пуассона. Теорема Пуассона для схемы Бернулли

Nota. Если вероятность успеха p в схеме Бернулли мала или близка к 1, то предельная формула Лапласа при недостаточно большом числе испытаний дает достаточно большую погрешность. В этой ситуации следует использовать формулу Пуасоона (формула редких событий)

Схема: вероятность числа успеха при одном испытании p_n зависит от числа испытаний n, причем таким образом, что $np_n \approx \lambda = const$

λ - интенсивность появления редких событий в единицу времени в потоке испытаний

Th. 1. (формула Пуассона) Пусть $n \to \infty, p_n \to 0$ таким образом, что $np_n \to \lambda = const > 0$

Тогда вероятность k успехов при n испытаниях: $P_n(k) = C_n^k p_n^k (1-p_n)^{n-k} \underset{n\to\infty}{\to} = \dfrac{\lambda^k}{k!}e^{-\lambda}$

□

Обозначим $\lambda_n = np_n$. Тогда $p_n = \dfrac{\lambda_n}{n}$ и

$$P_n(k) = C_n^k \left(\frac{\lambda_n}{n}\right)^k \left(1-\frac{\lambda_n}{n}\right)^{n-k} \underset{n\to\infty}{\to} \frac{n^k}{k!}\frac{\lambda_n^k}{n^k}\left(1-\frac{\lambda_n}{n}\right)^n \cancelto{1}{\left(1-\frac{\lambda_n}{n}\right)^{-k}} \underset{n\to\infty}{\to} \frac{\lambda_n^k}{k!}\left(1-\frac{\lambda_n}{n}\right) =$$

$$\frac{\lambda_n^k}{k!}\left(\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right)^{-\frac{\lambda_n}{n}n} \underset{n\to\infty}{\to} \frac{\lambda_n^k}{k!}e^{-\lambda_n} \underset{n\to\infty}{\to} \frac{\lambda^k}{k!}e^{-\lambda}$$

□

Th. 2. (оценка погрешности в формуле Пуассона) Пусть v_n - число успехов при n испытаниях в схеме Бернулли

p - вероятность успеха при одном испытании, $\lambda = np$, $A \subset \{0, 1, \dots, n\}$ - произвольное подмножество чисел

Тогда $|P_n(v_n \in A) - \sum\limits_{k\in A}\dfrac{\lambda^k}{k!}e^{-\lambda}| \leq \min(p, np^2) = \min(p, p\lambda)$

(без доказательства)

Def. Соответствие $k \to \dfrac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, \dots$ называется распределением Пуассона с параметром $\lambda > 0$ (обозначается Π_λ)

Ex. Прибор состоит из 1000 элементов, вероятность отказа каждого элемента равна 0.001. Какова вероятность отказа больше двух элементов

$P_n(k) \approx \dfrac{\lambda^k}{k!}e^{-\lambda}$

$n = 1000, p = 0.001, \lambda = 1$

$P_n(k > 2) = 1 - P_n(k \leq 2) = 1 - P(0) - P(1) - P(2) \approx 1 - \left(\frac{1^0}{0!}e^{-1} + \frac{1^1}{1!}e^{-1} + \frac{1^2}{2!}e^{-1}\right) = 1 - \left(1 + 1 + \frac{1}{2}\right)e^{-1} \approx$ 0.0803