

# Закон больших чисел (Law of large numbers)

Закон больших чисел в теории вероятности утверждает, что совместное действие набора случайных факторов может привести к неслучайному (детерминированному) результату. Этим принципом объясняется сближение частоты наступления случайного события с его вероятностью при возрастании числа испытаний.

Простейший пример — опыт с бросанием монеты. Теоретически выпадение орла или решки равновероятно. Это значит, что если подбросить монету 10 раз, то 5 раз должен выпасть орел и 5 раз — решка. Однако общеизвестно, что вероятность этого очень мала. С тем же успехом может выпасть 9 к 1, 3 к 5 и т.д. Тем не менее, если увеличить число опытов, скажем, до 100, то вероятность выпадения орла или решки приблизится к 50%. В пределе, если устремить число опытов к бесконечности, то вероятность выпадения орла и решки будет асимптотически стремиться к 50%.

То, какой стороной упадет монета, зависит от множества случайных факторов: как она будет лежать на ладони у экспериментатора, силы броска, высоты падения, скорости и т.д. Тем не менее, при достаточно большом числе опытов независимо от действия этих факторов мы всегда можем утверждать, что эмпирическая (опытная) вероятность будет близка к теоретической.

Например, пусть требуется оценить доходы населения в некотором регионе. Если мы рассмотрим 10 наблюдений, в которых у 9 респондентов доходы были около 20 000, а у одного — 500 000, то расчет простого среднего покажет доход на уровне 68 000, что не отражает реальную картину.

Если же мы рассмотрим 100 наблюдений, из которых 99 покажут доход 20 000 и только один — 500 000, то среднее составит около 24 800, что более адекватно отражает реальную ситуацию. При увеличении числа наблюдений среднее будет стремиться к своему истинному значению.

Именно закон больших чисел при анализе данных требует «набрать статистику», т.е. использовать как можно большее число наблюдений для получения достоверных результатов. При этом возникает противоречие между желаемой точностью результата и затратами на сбор и обработку данных.