Кидаємо монету
$2n$ разів — орел чи решка?
Яка ймовірність, що
буде m орлів та
$2n-m$ решок?
Будемо записувати
орли як 1 та решки
як 0. Отримаємо послідовність
нулів та одиниць, наприклад
1 0 0 1 1 1 0 1 $(n=4, m=5)$
Скільки є таких
послідовностей? 2^{2n}
А скільки між ними
таких, щоб було m
орлів? Ми мусимо
обрати m позицій

З $2n$ можливих, тобто отримуємо $\binom{2n}{m} = \frac{(2n)!}{m!(2n-m)!}$

Так що маємо

$p(m) = \frac{\binom{2n}{m}}{2^{2n}}$ (це носить назву розподіла Бернуллі)

Яка кількість орлів найбільш імовірна?

Звісно, n

Але можна показати, що ця імовірність для дуже великих n досить мала: $\sim \frac{1}{\sqrt{\pi n}}$

Але будь-який інший результат ще менш імовірний.

Наскільки менш імовірний?

Тобто ми хочемо оцінити частку $\frac{p(n+a)}{p(n)}$

(помітимо, що $p(n+a)=p(n-a)$).

Отримуємо:

$$\frac{p(n+a)}{p(n)} = \frac{\binom{2n}{n+a}}{\binom{2n}{n}} = \frac{n!^2}{(n+a)!\,(n-a)!}$$

$$= \frac{n(n-1)\cdots(n-a+1)}{(n+1)\cdot\ldots\cdot(n+a)} =$$

$$\frac{\left(1-\frac{1}{n}\right)\cdot\ldots\cdot\left(1-\frac{a-1}{n}\right)}{\left(1+\frac{1}{n}\right)\cdot\ldots\cdot\left(1+\frac{a}{n}\right)}$$

Добуток складніше оцінювати, ніж суму, так що давайте прологарифмуємо:

$$\log\frac{p(n+a)}{p(n)} = \left(\log\left(1-\frac{1}{n}\right)+\cdots+\log\left(1-\frac{a-1}{n}\right)\right)$$

$$-\left(\log\left(1+\frac{1}{n}\right)+\cdots+\log\left(1+\frac{a}{n}\right)\right)$$

Нас цікавить випадок, коли $a \ll n$, інакше частка $\frac{p(n+a)}{p(n)}$ дуже мала.

У цьому випадку, ми можемо застосувати лінійне наближення

$$\log(1+y) \approx y \quad , \quad |y| \ll 1.$$

(це не складно довести, але в нас немає часу).

Тоді отримуємо:

$$\log \frac{p(n+a)}{p(n)} \approx -\left(\frac{1}{n} + \dots + \frac{a-1}{n}\right) - \left(\frac{1}{n} + \dots + \frac{a}{n}\right) = -\left(\frac{a(a-1)}{2n} + \frac{a(a+1)}{2n}\right)$$

$$= -\frac{a^2}{n}.$$

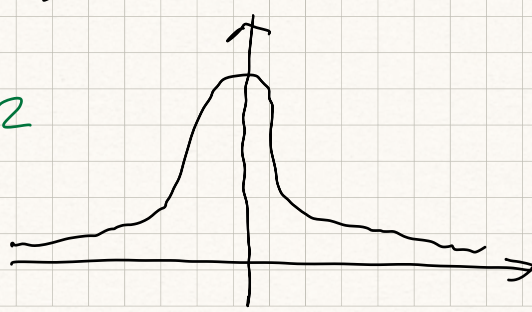
Тобто якщо $a \approx x\sqrt{n}$,

маємо $\log \frac{p(n+a)}{p(n)} \sim -x^2$, або

$$\frac{p(n+a)}{p(n)} \approx e^{-x^2}.$$

Ми довели найпростіший випадок центральної граничної теореми, яка стверджує, що коли $n \to \infty$, розподіл Бернуллі збігається до нормального, чи Гауссова розподілу, гістограма якого задається функцією

$$p(x) = C \cdot e^{-x^2}$$



де C є якесь позитивне число.

Щоб знайти C, згадаємо що сума ймовірностей усіх можливих варіантів повинна рівнятися 1. Тобто маємо

$$\int_{-\infty}^{\infty} Ce^{-x^2}\,dx = 1,$$

так що

$$C = \frac{1}{\int_{-\infty}^{\infty} e^{-x^2}\,dx}$$

Залишається обчислити цей інтеграл. Нажаль, невизначений інтеграл

$\int_{-\infty}^{y} e^{-x^2}dx$ не обчислюється в елементарних функціях (це так звана "error function" erf). Але визначений (невласний) інтеграл можна обчислити, якщо застосувати інтегрування функцій двох змінних та полярні координати:

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2} dx\,dy$$

$$= \int_{\mathbb{R}^2} e^{-x^2-y^2} dx\,dy = \begin{pmatrix}\text{перехід до} \\ \text{полярних} \\ \text{координат}\end{pmatrix}$$

$$\int_0^{\infty} r\,dr \int_0^{2\pi} d\theta \cdot e^{-r^2} = 2\pi \int_0^{\infty} e^{-r^2} r\,dr =$$

$$\boxed{s = r^2} = 2\pi \int_0^{\infty} e^{-s} \frac{ds}{2} =$$

$$= \pi \int_0^\infty e^{-s}\, ds = -\pi e^{-t} \Big|_0^\infty = \pi .$$

Тобто $\displaystyle\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

Так що

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

Це дозволяє обчислити шанс, що

$$-b\sqrt{n} \le a \le b\sqrt{n}$$

для великих n.

Це

$$p = \frac{1}{\sqrt{\pi}} \int_{-b}^{b} e^{-x^2} dx .$$