1. Experimento probabilístico

El láser emite N fotones cada segundo. Cada fotón tiene una energía desconocida ϵ. Si no se utiliza el cristal semitransparente, el sensor medirá una energía de $E_{max} = N\epsilon$ durante un segundo, pero esto no nos permitirá determinar N.
Por otro lado, si se interpone el cristal, cada fotón tiene una probabilidad $p = 1/2$ de atravesarlo, lo que significa que la energía medida por el sensor en un segundo fluctuará. La probabilidad de recibir n fotones en un segundo viene dada entonces por la fórmula binomial

$$P(n,N) = \frac{N!}{n!(N-n)!}\left(\frac{1}{2}\right)^N$$

Si realizamos un gran número de mediciones de un segundo con nuestro sensor, la cantidad de energía recibida $E = n\epsilon$ será diferente en cada medición, y veremos que las mediciones se distribuyen según una distribución que tiende a una gaussiana. Podemos entonces medir el valor medio $\langle E\rangle$ y la desviación típica ΔE. Ahora

$$\langle E\rangle = \langle n\rangle\epsilon$$
$$(\Delta E)^2 = \left\langle E^2\right\rangle - \langle E\rangle^2 = \left(\left\langle n^2\right\rangle - \langle n\rangle^2\right)\epsilon^2$$

Y sabemos que para una distribución binomial

$$\langle n\rangle = Np = N/2$$
$$\left\langle n^2\right\rangle - \langle n\rangle^2 = Np(1-p) = N/4$$

así que

$$\langle E\rangle = N\epsilon/2$$
$$(\Delta E)^2 = N\epsilon^2/4$$
$$\frac{\Delta E}{\langle E\rangle} = \frac{1}{\sqrt{N}}$$

Las fluctuaciones de energía nos permiten calcular N: si N fuera infinito (en otras palabras, si la luz fuera un flujo continuo), siempre recibiríamos exactamente $E = E_{\max}/2$ detrás del cristal. Las fluctuaciones son, por tanto, la huella de la naturaleza discreta de un fenómeno, y nos permiten calcular el número de partículas elementales implicadas aunque no conozcamos sus propiedades microscópicas, como ϵ.

2. Control de calidad

Por término medio, cada motor se avería una vez al mes, es decir, cada 720 horas. Por consiguiente, la probabilidad de que un motor determinado falle durante una hora es de $1/720$. La probabilidad de este fallo durante un vuelo de 5 horas es, por tanto, de

$$p = 5/720 = 1/144$$

Si un avión tiene un total de n motores, la probabilidad de que k de sus motores fallen durante un solo vuelo viene dada por la fórmula binomial

$$P(k,n) = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

Habrá un accidente si $k = n-1$ o $k = n$, es decir, si quedan 1 ó 0 motores válidos. Esto da la probabilidad de que se produzca un accidente si hay n motores.

$$\begin{aligned} P(\text{caída}, n) &= P(n-1, n) + P(n, n) \\ &= \frac{n!}{(n-1)!} p^{n-1}(1-p) + \frac{n!}{n!} p^n \\ &= p^{n-1}(n(1-p) + p) \end{aligned}$$

Ahora podemos ver los primeros valores de n :

$$\begin{aligned} P(\text{caída}, 1) &= 1 \\ P(\text{caída}, 2) &= 1{,}38\ 10^{-2} \\ P(\text{caída}, 3) &= 1{,}44\ 10^{-4} \\ P(\text{caída}, 4) &= 1{,}33\ 10^{-6} \end{aligned}$$

Si sólo tenemos 3 motores, tenemos una media de $10000 \times P(\text{caída}, 3) = 1{,}44$ accidentes cada 10000 vuelos, así que necesitamos 4 motores, en cuyo caso tenemos $1{,}33$ accidentes cada millón de vuelos.

3. Red eléctrica

En primer lugar, calculamos el número de tuberías. Cada nodo tiene 4 tuberías, y cada tubería es compartida por 2 vecinos, por lo que hay

$$N = 4 \times M/2 = 2M$$

tuberías (en realidad $2M - 2\sqrt{M}$ si se tienen en cuenta los bordes). Por tanto, el número total de configuraciones para n tuberías bloqueadas de N es

$$\Omega_{tot} = \frac{N!}{n!(N-n)!}$$

Ahora queremos contar el número de configuraciones en las que hay cuatro tuberías bloqueadas que van al mismo nodo. La primera tubería tiene $N - (n-1)$ lugares para ir sin tapar otra tubería bloqueada, y entonces se determina la posición de las otras tres (de lo contrario se cuenta varias veces la misma configuración). Quedan $n-4$ tuberías para colocar libremente de entre $N-4$, por lo que en total

$$\Omega(1) = (N-n+1) C_{N-4}^{n-4} = (N-n+1) \frac{(N-4)!}{(n-4)!(N-n)!}$$

Por tanto, la probabilidad de que haya al menos un nodo taponado es

$$\begin{aligned} P(1) = \frac{\Omega(1)}{\Omega_{tot}} &= (N-n+1) \frac{(N-4)!}{N!} \frac{n!}{(n-4)!} \\ &= (N-n+1) \frac{n(n-1)(n-2)(n-3)}{N(N-1)(N-2)(N-3)} \end{aligned}$$

Nota 1: existe exactamente la misma probabilidad para cualquier otra forma formada por exactamente 4 tubos, por ejemplo poniéndolos en un cuadrado o en línea, o incluso a siete nodos de distancia. Pero en todos estos otros casos, ningún nodo está completamente seco, así que no importa.
Nota 2: Puede haber situaciones más complicadas en las que ningún nodo individual esté seco, pero las tuberías que rodean a un conjunto de nodos (un barrio entero) estén bloqueadas. Esto es mucho menos probable si $n \ll N$, pero ocurre. Si queremos incluir estos casos, el cálculo se complica enormemente, hasta el punto de que no hay solución exacta para todos los N.

4. Secuenciación genética

Cálculo aproximado El procedimiento más intuitivo no es exacto, pero funciona bastante bien cuando k no es demasiado pequeño en comparación con n. Consiste en el siguiente razonamiento: sea cual sea el contenido de la primera hebra, contiene $n-k+1$ secuencias (A.N. 5 secuencias de longitud 4 de 8 nucleótidos). También hay $n-k+1$ en la hebra opuesta, es decir, un total de $(n-k+1)^2$ pares formados por una secuencia de cada hebra.

$$\text{Hebra 1}: \underbrace{ATAG}_{\text{Secuencia 1}}\ CCAT$$
$$\text{Hebra 2}: GC\ \underbrace{AGCG}_{\text{Secuencia 3}}\ AA$$

La probabilidad de que dos secuencias de longitud k sean exactamente idénticas es la probabilidad de que el primer nucleótido sea el mismo (1/4) Y de que el segundo sea idéntico (1/4) etc. es decir $(1/4)^k$. Si multiplicamos esta probabilidad por el número de pares de secuencias, obtenemos el número medio de secuencias idénticas entre las dos cadenas:

$$\lambda = (n-k+1)^2 \frac{1}{4^k}$$

A.N. $\lambda = 5^2/4^4 \approx 0.1$.
Visto de otra forma, hay $N = (n-k+1)$ secuencias en la primera hebra, y cada una tiene una probabilidad $p = (n-k+1)/4^k$ de tener un gemelo opuesto. Si $p \ll 1$ podemos utilizar una distribución de Poisson con $\lambda = Np$. Esto da la probabilidad de que haya **exactamente** un par de secuencias idénticas en las dos cadenas:

$$P(1,\lambda) = \lambda e^{-\lambda}$$

y la probabilidad de que haya **al menos** un par de secuencias idénticas :

$$1 - P(0,\lambda) = 1 - e^{-\lambda}$$

Cuando λ es suficientemente pequeño, $e^{-\lambda} \approx 1-\lambda$ y por lo tanto

$$P(1,\lambda) \approx 1 - P(0,\lambda) \approx \lambda$$

A.N. $P(1) \approx 9,1\%, 1 - P(0) \approx 9,3\%$.
Problema: con este cálculo, si n es suficientemente grande, podemos tener una probabilidad $p = (n-k+1)/4^k > 1$, lo cual es obviamente absurdo.

Límites de la aproximación intuitiva El cálculo anterior es inexacto por una sencilla razón: las $n-k+1$ secuencias contenidas en una sola hebra no son necesariamente todas diferentes - de hecho, si $n-k+1 > 4^k$ es muy probable que haya varios casos de la misma secuencia. Sin embargo, si hay dos secuencias idénticas en la misma hebra, ésta contiene una menor diversidad de secuencias y es menos probable que tenga una gemela en otra hebra.
El cálculo intuitivo funciona bien si $n-k+1 \ll 4^k$ (hebra suficientemente corta y secuencias suficientemente largas, de modo que hay pocas probabilidades de encontrar la misma secuencia varias veces). En el otro límite (hebra muy larga y secuencias bastante cortas) tenemos que sustituir el número total de secuencias $N = n-k+1$ por el número medio de secuencias diferentes $\langle D \rangle$ en una hebra. Si $N > 4^k$ entonces necesariamente

$$D = 4^k$$

y entonces la probabilidad de tener un gemelo es $p = D/4^k = 1$ (las dos cadenas contienen todas las secuencias posibles). Por otra parte, si $N < 4^k$, entonces la probabilidad de que D de estas N secuencias sean diferentes es la misma que la del problema del cumpleaños.

$$\frac{4^k \left(4^k-1\right)\dots\left(4^k-D+1\right)}{\left(4^k\right)^D}$$

(las secuencias son cortas, por lo que la probabilidad de que se solapen es baja, y pueden considerarse independientes) y el valor medio de D viene dado por

$$\langle D\rangle = \sum_{D=1}^{N} D \frac{\left(4^k\right)!}{(4^k-D)!4^{kD}}$$

Este cálculo es bastante más difícil, por lo que nos limitaremos a la primera versión.

Nota sobre el problema problema del cumpleaños:
¿Cuál es la probabilidad de que dos personas cumplan años el mismo día en una clase de $N = 23$ personas?
La probabilidad P de que haya al menos dos personas con el mismo cumpleaños en una clase de N personas viene dada por $1 - P_0$, siendo P_0 la probabilidad de que todos los cumpleaños sean diferentes. Para ello, el primer alumno tiene 365 días posibles, el segundo sólo 364, etc., mientras que hay 365^N posibilidades para los cumpleaños de todos los alumnos, por lo que

$$P_0 = \frac{365 \times 364 \times \cdots (365 - N + 1)}{365^N}$$

Si $N = 23$, $P_0 \simeq 0.497$ lo que implica $P = 1 - P_0 = 50.3\%$