

Física Biológica/Biofísica

Tercera Teórica

La clase anterior vimos conceptos relacionados a combinatoria y probabilidades, probabilidad condicional, variables aleatorias y terminamos hablando de distribuciones de probabilidad.

Dada una variable aleatoria, X , real, la distribución de probabilidad es una función que le asigna un valor mayor o igual a cero a los valores que puede tomar X .

Las variables aleatorias pueden ser discretas (casi todos los ejemplos que vimos la clase pasada eran así) o continuas.

En el caso de variables discretas, donde X toma uno de N valores posibles, x_1, \dots, x_N , la distribución de probabilidad asigna un valor $P(x_i) \geq 0$ a cada x_i tq la suma sobre i de $P(x_i)$ es 1.

En el caso de una variable continua se define la densidad de probabilidad, \mathcal{P} , de forma que $\mathcal{P}(x) dx$ es la probabilidad de que la variable aleatoria, X , tome un valor entre x y $x + dx$. $\mathcal{P}(x)$ puede ser mayor que 1. La integral de $\mathcal{P}(x)$ sobre todos los valores posibles de x es 1.

La probabilidad es adimensional, la densidad de probabilidad no. En el caso de una variable continua:

$\mathcal{P}(x) dx$ es adimensional

$dP = P(x) dx$ es una probabilidad diferencial (la probabilidad que le asigno a un intervalo diferencial de valores posibles de la variable aleatoria). dP es adimensional, \mathcal{P} tiene unidades de $1/[x]$

Es lo mismo que la masa y la densidad de masa. Si tengo una cuerda a la que subdivido en pedacitos diferenciales de longitud dx y a cada uno le asigno un diferencial de masa, dm , puedo definir la densidad lineal de masa, ρ , tal que $dm = \rho dx$. La densidad, ρ , tiene unidades de masa por unidad de longitud mientras que dm tiene unidades de masa. Lo mismo pasa con \mathcal{P} y dP (con dP adimensional por lo que si $[x]=\text{cm}$, entonces \mathcal{P} tiene unidades de $1/\text{cm}$).

En el caso de variables continuas es común trabajar con la distribución de probabilidad acumulada, en lugar de con la densidad de probabilidad. La distribución acumulada también se define para variables discretas.

Por otro lado, dada la distribución de probabilidad de una variable aleatoria, X , se definen una serie de cantidades que sirven para caracterizarla:

- Soporte
- Valor medio ($\langle X \rangle$) o valor esperado ($E(X)$)
- Moda
- Mediana
- Cuantiles
- Varianza
- Desvío standard
- Skewness
- Kurtosis

Pasemos al “pizarrón” para definir todas estas cosas.

Veamos también qué pasa cuando se tiene más de una variable aleatoria, X_1, \dots, X_k y qué cantidades se pueden definir en ese caso (e.g., covarianza).

Dos resultados importantes

Ley de los grandes números.

El promedio de n variables aleatorias, X_1, \dots, X_n , equidistribuidas, (con $E[X_i] = \langle X_i \rangle = \mu$, finita) e independientes entre sí

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n} \quad \text{tiende a } \mu \text{ con probabilidad 1.} \quad P\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1,$$

Teorema central del límite.

La suma de n variables aleatorias, X_1, \dots, X_n , independientes entre sí:

$$S_n := X_1 + \dots + X_n = \sum_{i=1}^n X_i \quad \text{donde } E[X_i] = \mu \text{ y } \text{Var}[X_i] = \sigma^2 \quad \text{finitas}$$

es una variable aleatoria cuya distribución converge a la $\mathcal{N}(n\mu, n\sigma^2)$ para $n \rightarrow \infty$. Equivalente a decir que la distribución del promedio:

$$\bar{X}_n \equiv \frac{X_1 + \dots + X_n}{n} \quad \text{converge a } \mathcal{N}(\mu, \sigma^2/n) \text{ (i.e., desvío: } \sigma/n^{1/2}\text{).}$$

Algunas distribuciones “famosas”. Variables discretas

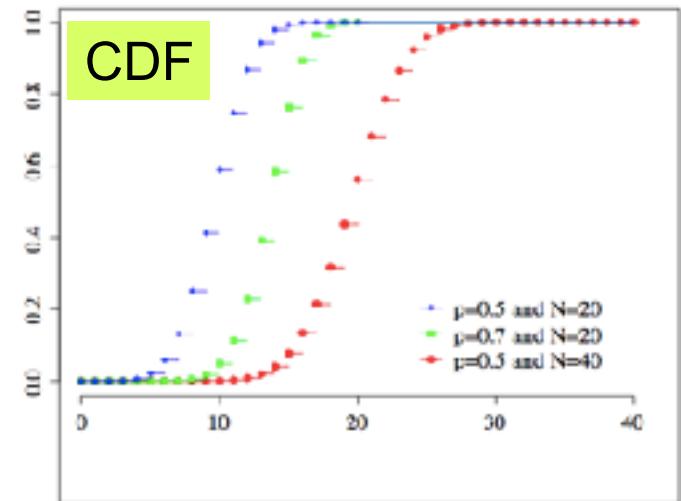
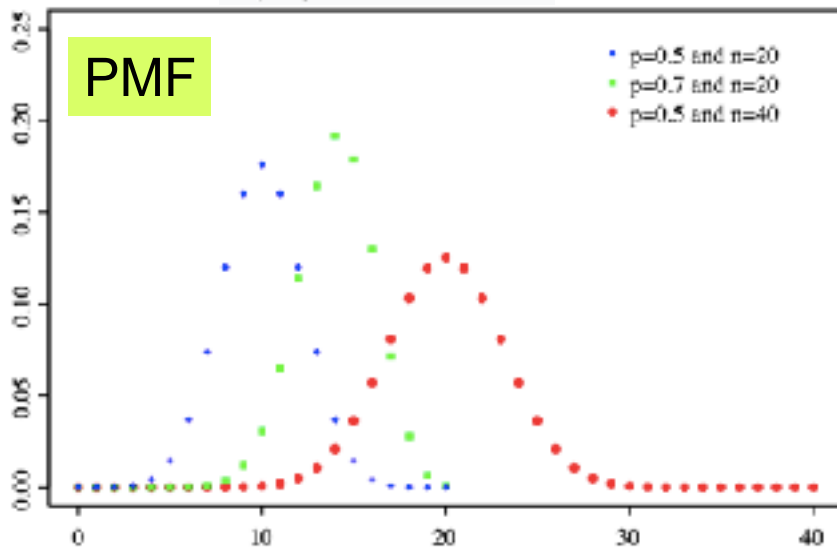
Distribución binomial, $B(n,p)$

Dada una secuencia de n experimentos cada uno de los cuales puede dar dos resultados posibles ($X=1$ con probabilidad, p , y $X=0$ con $q=1-p$), la distribución binomial es la que me dice cuál es la probabilidad, $P(k)$, de que $X=1$ haya sucedido k veces entre los n . O sea, corresponde a la variable aleatoria: número de veces que sale 1.

$$P(k) = \binom{n}{k} p^k q^{n-k}$$

$$\langle k \rangle = np$$

$$\text{Var} = np(1-p)$$



Algunas distribuciones “famosas”. Variables discretas

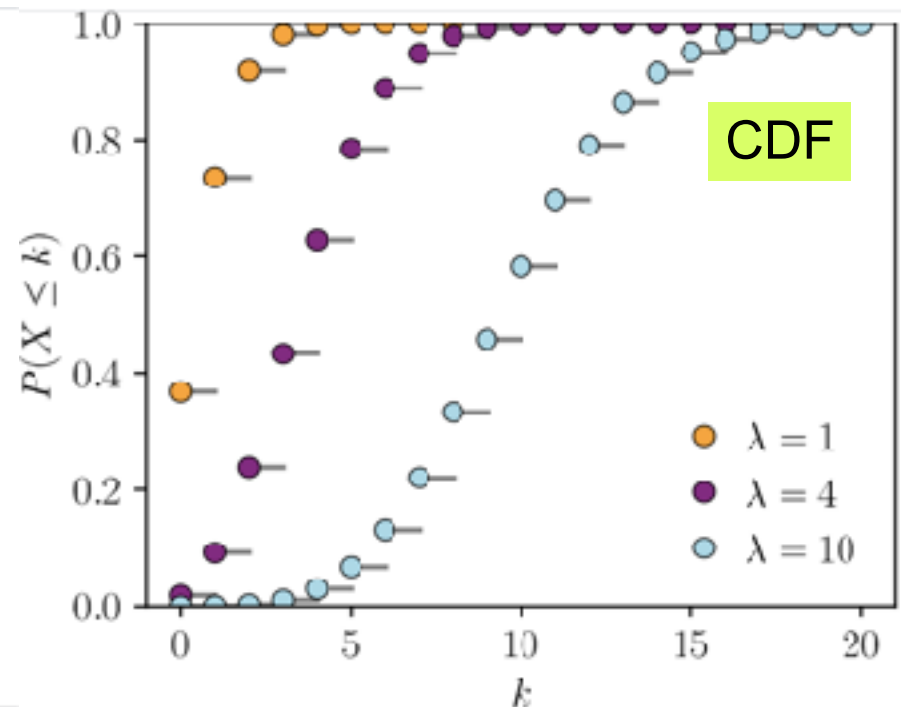
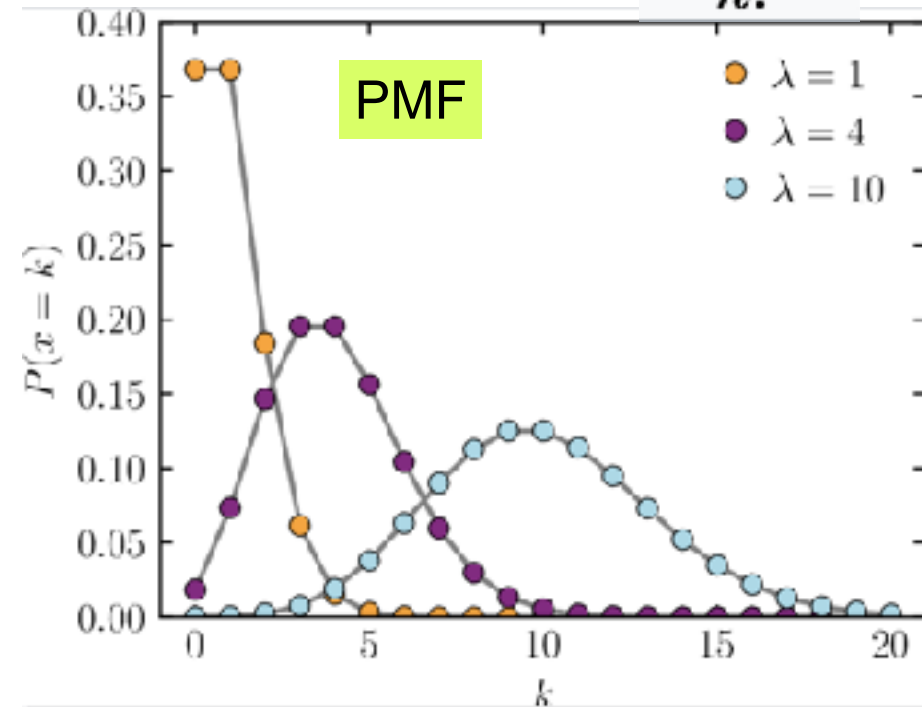
Distribución de Poisson, $\text{Pois}(\lambda)$

Se obtiene a partir de la $B(n,p)$ en el límite en que $n \rightarrow \infty$ y $p \rightarrow 0$, pero $pn \rightarrow \lambda$, finito. Es LA distribución discreta (la que describe cosas numerables, e.g., número de moléculas en un volumen, etc).

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\langle k \rangle = \lambda$$

$$\text{Var} = \lambda$$



Algunas distribuciones “famosas”. Variables discretas

Distribución multinomial: describe varias variables a la vez.

Parecida a la binomial, pero cuando los experimentos de la secuencia pueden tener k resultados distintos con probabilidades p_1, \dots, p_k ($p_1 + \dots + p_k = 1$). Dada una secuencia de n experimentos, la distribución multinomial da la probabilidad, $P(x_1, \dots, x_k)$, de que el i -ésimo resultado haya salido x_i veces (por lo tanto, $x_1 + \dots + x_k = n$). Es decir, describe k variables aleatorias: número de veces que sale x_1 , número de veces que sale x_2 , ..., número de veces que sale x_k .

$$P(x_1, \dots, x_k) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

$$\langle x_i \rangle = np_i \quad \text{Var}(x_i) = np_i(1-p_i)$$

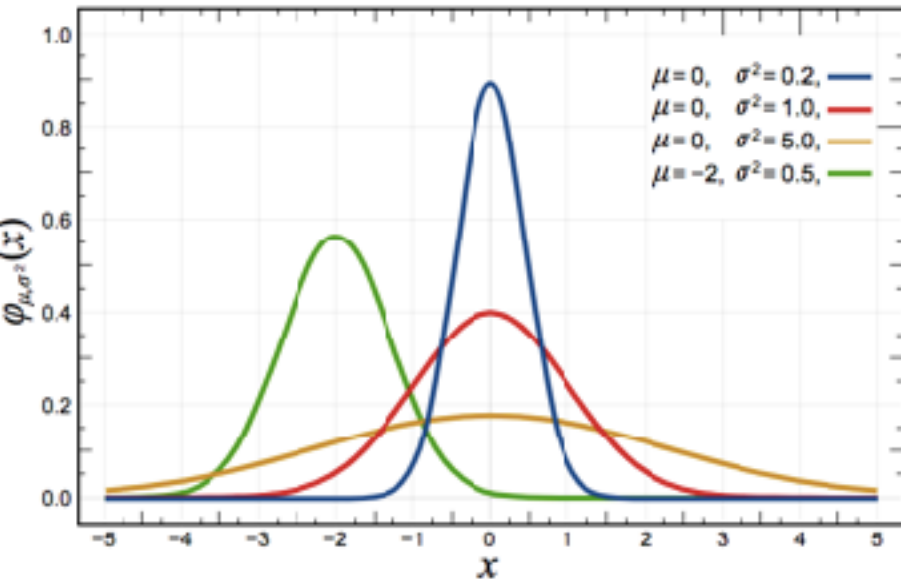
$$\text{Cov}(x_j, x_i) = -np_i p_j \quad (i \neq j)$$

Algunas distribuciones “famosas”. Variables continuas.

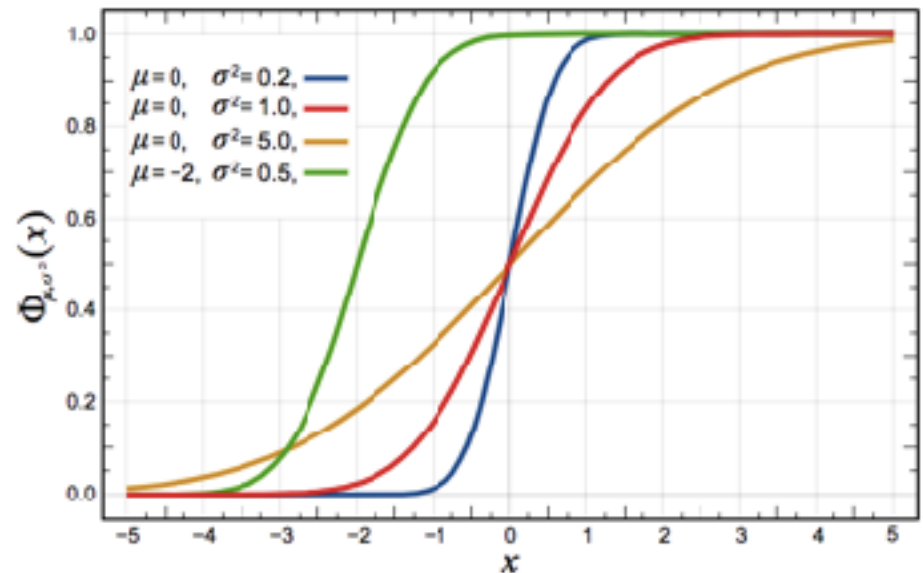
Distribución normal o Gaussiana

$$\langle x \rangle = \mu; \text{Var}(x) = \sigma^2$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

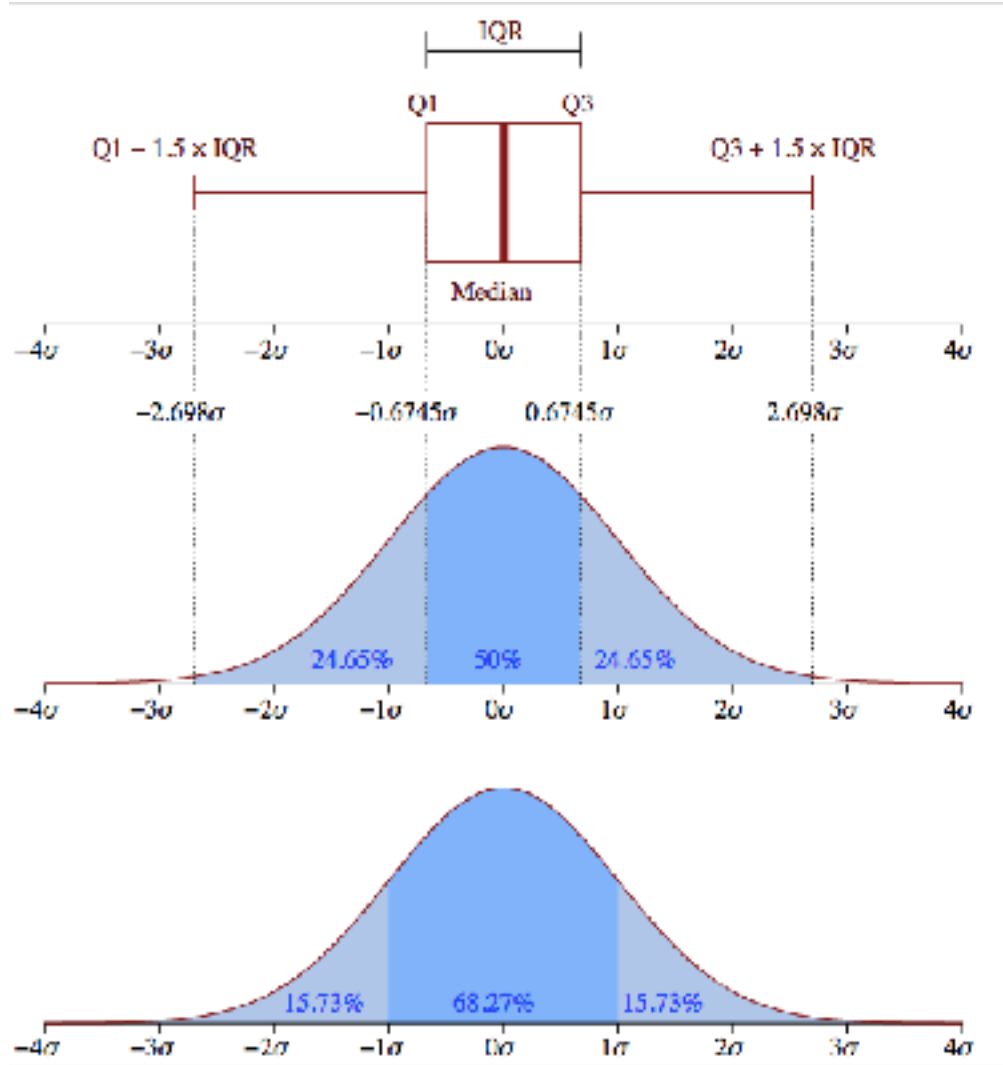


PDF



CDF

Distribución normal o Gaussiana, cuantiles.



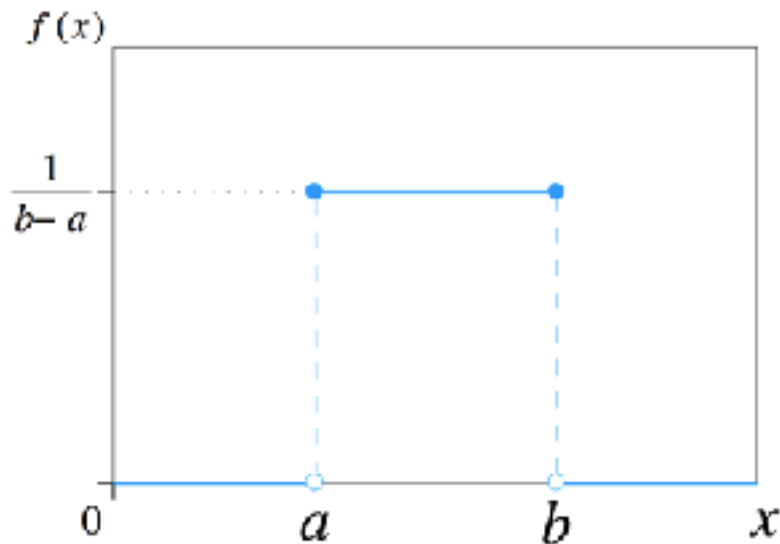
De Jhguch at en.wikipedia, CC BY-SA 2.5, <https://commons.wikimedia.org/w/index.php?curid=14524285>

Algunas distribuciones “famosas”. Variables continuas.

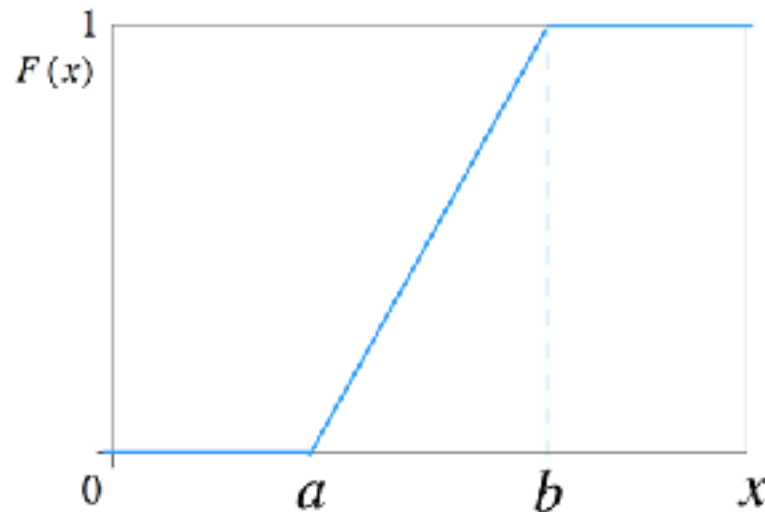
Distribución uniforme sobre $[a,b]$

$$\langle x \rangle = (b+a)/2 ; \text{Var}(x) = (b-a)^2/12$$

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$



PDF



CDF

Los generadores de números al azar generan números con distribución uniforme entre 0 y 1. Después se elige una función de los números generados que permitan pasar de una variable distribuida uniformemente entre 0 y 1 a otra con la distribución deseada (siempre que se pueda!). Esto se usa también para elegir variables aleatorias discretas. Lo vamos a ver en la práctica.

¿Por qué nos interesan tanto las probabilidades en esta materia?

Temperatura, ¿qué es?

Ley de los gases ideales

$$PV = N k_B T \quad \text{o} \quad PV = N_M RT$$

N =número de moléculas. N_M =número de moles P =presión,
 V =volumen, T =temperatura absoluta (medida en °K)

k_B =constante de Boltzmann= $R/N_A = 1.3806488 \times 10^{-23} \text{ J } \text{ }^\circ\text{K}^{-1}$

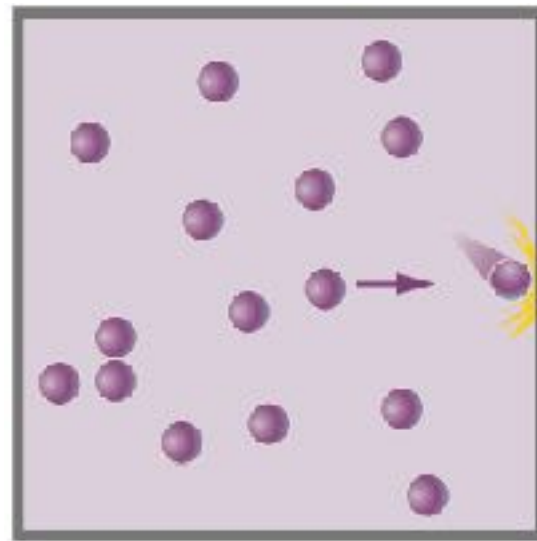
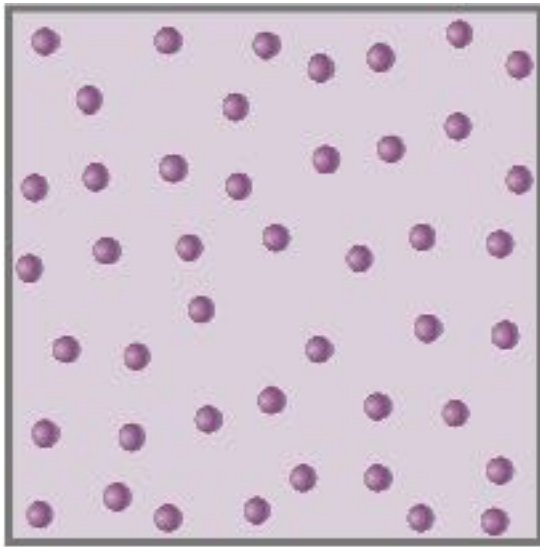
R = constante de los gases= $8.314 \text{ J}/(\text{mol } \text{ }^\circ\text{K})$

N_A =número de Avogadro= $6.02214179 \times 10^{23}/\text{mol}$

¿qué unidades tiene esta expresión?

Energía

Interpretación microscópica de esta ley.



Moléculas muy pequeñas comparadas con la distancia entre ellas que se mueven libremente (solo chocan elásticamente con las paredes del envase que las contiene). ¿Con qué velocidad se mueven?

Con una velocidad cualquiera, o sea, podemos pensar a la velocidad como una variable aleatoria continua. Me puedo preguntar entonces cuál es la distribución de esta variable (disclaimer: si no intercambiaran energía con nada de verdad, en realidad la distribución sería la de las condiciones iniciales y qué decir!).

La velocidad es un vector de 3 componentes

$$\mathbf{v} = v_x \hat{i} + v_y \hat{j} + v_z \hat{k}$$

caracterizado por 3 variables aleatorias, v_x , v_y , v_z , independientes entre sí. Pensemos en una de ellas, v_x , porque a todas les va a pasar lo mismo. Trabajemos como si viviéramos en 1 dimensión.

Como las moléculas son todas iguales y no interactúan entre sí puedo pensar que el conjunto de todas las moléculas del gas representa muchas realizaciones del mismo experimento realizado con una sola molécula. O sea, la fracción de moléculas que tienen velocidad entre v_x y $v_x + dv_x$ determina la densidad de probabilidad de v_x (también llamada función de distribución de la velocidad, $f(v_x)$). Así

$$dN(v_x)/N = f(v_x) dv_x$$

es la probabilidad de que una molécula tenga velocidad entre v_x y $v_x + dv_x$

En promedio hay tantas moléculas que se mueven con una dada v_x , como las que se mueven con $-v_x$. O sea,

$$\langle v_x \rangle = 0$$

¿cuál es la varianza alrededor de este valor medio?

Var $\sim T$

Sabemos que $k_B T$ tiene unidades de energía

Por otro lado, la molécula de un gas ideal, que no interactúa con nada, solo tiene energía cinética. En un mundo unidimensional, la energía cinética es $m v_x^2/2$, donde m es la masa de la molécula.

La función de distribución de velocidades (la densidad de probabilidad) es (en el mundo unidimensional):

$$f(v_x) dv_x = \left(\frac{m}{2\pi kT} \right)^{1/2} e^{-\frac{mv_x^2}{2kT}} dv_x; \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

una distribución Gaussiana (la distribución de Boltzmann) con valor medio 0 y $\sigma^2 = k_B T/m$. Es decir, la temperatura me dice dentro de qué rango, alrededor del valor medio, varía la velocidad.

Pero además, $\sigma^2 = \langle v_x^2 \rangle - \langle v_x \rangle^2 = k_B T/m$, por lo que:

$$m \langle v_x^2 \rangle / 2 = k_B T / 2$$

Y lo mismo pasa en cualquiera de las 3 direcciones espaciales:

$$m \langle v_x^2 \rangle / 2 = m \langle v_y^2 \rangle / 2 = m \langle v_z^2 \rangle / 2 = k_B T / 2; \text{ equipartición de } E$$

Volviendo al mundo 3D, como lo que pasa en una dirección es independiente de lo que pasa en otra, la densidad de probabilidad conjunta, $f(v_x, v_y, v_z)$ es el producto de las individuales:

$$f(v_x, v_y, v_z) = f(v_x)f(v_y)f(v_z)$$

que se puede reescribir como:

$$f(v) d^3 v = \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{mv^2}{2kT}} d^3 v,$$

donde:

$$v^2 = v_x^2 + v_y^2 + v_z^2$$

$$d^3 v = dv_x dv_y dv_z$$

La distribución de velocidades de las moléculas de soluto en soluciones diluidas también se puede describir con una distribución gaussiana de valor medio 0 y varianza proporcional a T.

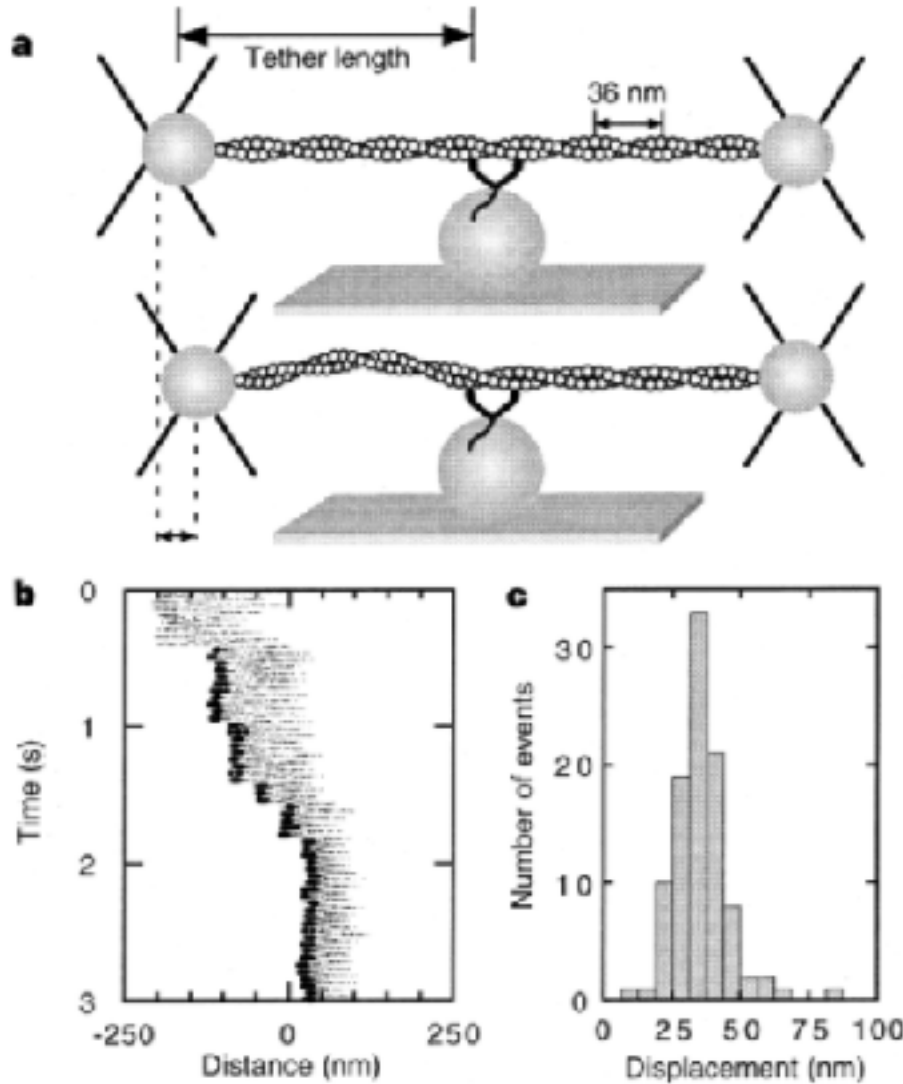
Esto se puede “derivar” del teorema central del límite considerando que el momento lineal de una molécula de soluto en un instante, t , es la suma del valor en un instante anterior, 0 , más la suma de todas las variaciones que sufrió debido a los choques con las moléculas de solvente entre 0 y t . Suponiendo que cada choque le da una variación independiente de la anterior (tomada siempre de la misma distribución de probabilidad que tiene que ver con el estado del solvente) y que es igualmente probable que sea en un valor mv_x que en $-mv_x$, esa suma va a tender a estar distribuida de forma normal alrededor de 0. El ancho de la normal va a determinar el valor de $\langle v_x \rangle^2$ y, por lo tanto, de la temperatura.

¿A qué energía característica corresponde $k_B T$?

A T ambiente (25C) es $1.38 \cdot 10^{-23} \cdot 298 \text{ J} = 4.11 \cdot 10^{-21} \text{ N m} \sim 4 \text{ pN nm}$

Comparemos este valor con los de fuerza y desplazamiento medidos en motores moleculares (como los involucrados en contracción muscular)

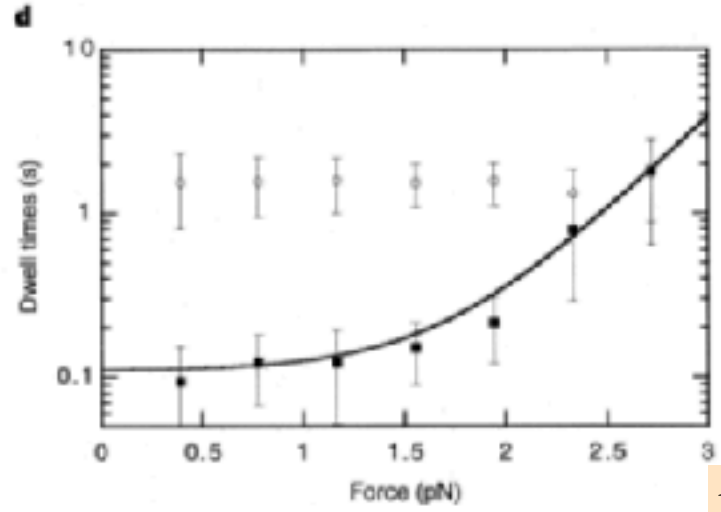
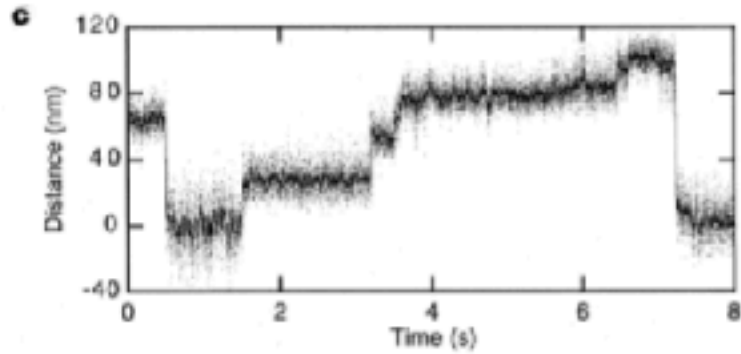
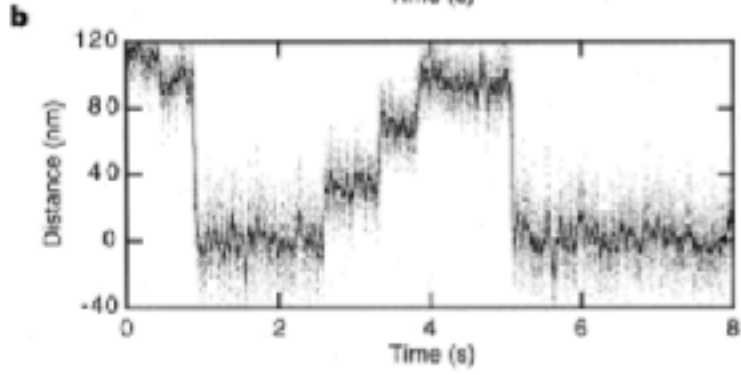
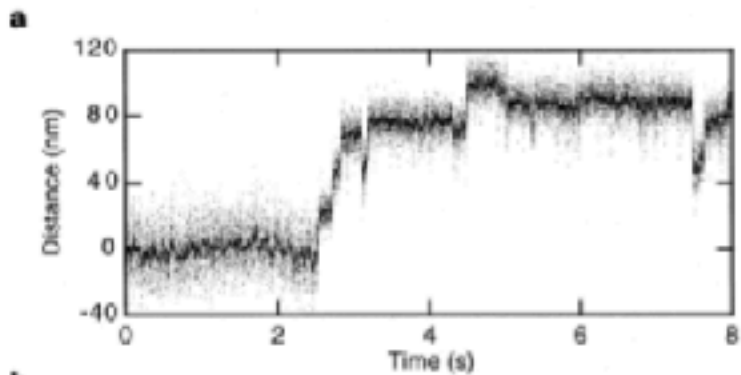
Myosin-V is a processive actin-based motor; Mehta et al, Nature 1999



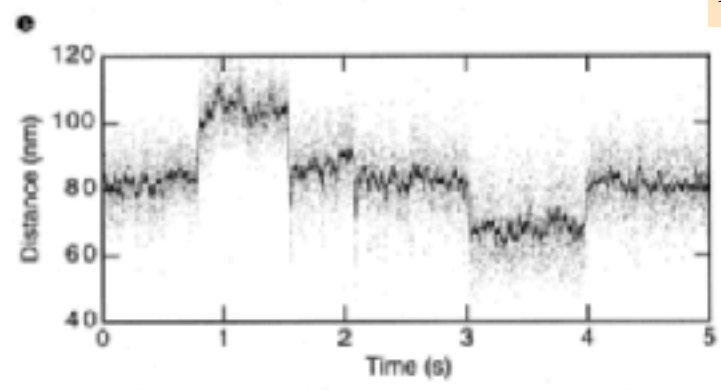
Se pegan cuentas (beads) a los extremos de un filamento de actina y se usan pinzas ópticas para colocarla sobre otra que está fija y cubierta de fragmentos de miosina. Las pinzas ópticas deben ejercer una fuerza que se opone a la de la interacción entre la miosina y la actina para mantener el equilibrio. Así se puede medir la fuerza.

Step measurements. a, The scheme of the oscillation experiments (see Methods). b, Sample trace at 2 mM ATP. The myosin molecule bound the actin just before the 0.5-s mark. After this, the bead ceased to follow the driving triangle wave for a fraction of each oscillation. c, Sample histogram of step increments so measured.

Myosin-V is a processive actin-based motor; Mehta et al, Nature 1999



$k_B T \sim 4 \text{ pN nm}$



Las fluctuaciones no son nada despreciables!

Processive stepping by myosin-V observed using the dual-beam optical trap. a-c, Sample traces of stepping behaviour at 2 mM, 10 mM and 1 mM ATP, respectively. Dwell periods T_1 at 1 μ M ATP fit single exponential statistics (circles). d, Mean dwell time as a function of load, at 2 mM ATP (squares) and at 1 mM ATP (circles). e, Backward stepping behaviour (at 1.5 s and 3 s) at high load in 1 mM ATP.