Capítulo 3

Distribuciones de probabilidad multivariadas

Sobre un dado espacio muestral podemos definir diferentes variables aleatorias. Por ejemplo, en un experimento binomial, X_1 podría ser la variable binomial (número de total de éxitos) y X_2 el número de éxitos en las k primeras pruebas. Si las variables aleatorias $X_1, X_2, \ldots, X_n$ están definidas sobre el mismo espacio muestral S se dice que están conjuntamente distribuidas.

3.1. Distribución de probabilidad conjunta

La función de distribución **conjunta** para las variables aleatorias $X_1, X_2, \ldots, X_n$ se define como

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \equiv \mathrm{Prob}\,\{X_1 < x_1,\ldots,X_n < x_n\} \tag{3.1}$$

donde $\{X_1 < x_1,\ldots,X_n < x_n\} = \{X_1 < x_1\} \cap \{X_1 < x_2\} \cap \ldots \cap \{X_n < x_n\}$. En otras palabras, $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ es la probabilidad de que las variables aleatorias X_i tomen *simultaneamente* valores en los intervalos $\{-\infty < X_i < x_i\}$ con $i = 1,\ldots,n$. La **densidad de probabilidad conjunta** $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ se define entonces como

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \frac{\partial^n F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{\partial x_1 \ldots \partial x_n} \tag{3.2}$$

de tal manera que

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} dx_1' \ldots dx_n' \; f_{X_1,\ldots,X_n}(x_1',\ldots,x_n') \tag{3.3}$$

Consideremos por simplicidad de aqui en mas solo dos variables aleatorias X e Y, con funciones de distribución y densidad $F_{X,Y}(x,y)$ y $f_{X,Y}(x,y)$ respectivamente. La función de distribución satisface $F_{X,Y}(-\infty,y) = F_{X,Y}(x,-\infty) = F_{X,Y}(-\infty,-\infty) = 0$ y $F_{X,Y}(\infty,\infty) = 1$. Ademas:

$$F_{X,Y}(x_2,y) - F_{X,Y}(x_1,y) = \mathrm{Prob}\,\{x_1 < X_1 < x_2; Y < y\}$$

La densidad de probabilidad $f_{X,Y}(x,y)$ satisface las condiciones $f_{X,Y}(x,y) \geq 0$ y

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; f_{X,Y}(x,y) = 1 \tag{3.4}$$

Dadas dos variables X e Y distribuidas conjuntamente, la función de distribución **reducida** o **marginal** $F_X(x)$ para la variable X viene dada por

$$F_X(x) = F_{X,Y}(x,\infty) = \int_{-\infty}^{x} dx' \int_{-\infty}^{\infty} dy \; f_{X,Y}(x',y) \tag{3.5}$$

Y de la misma manera, la densidad de probabilidad marginal $f_X(x)$ para la variable X se define como

$$f_X(x) = \int_{-\infty}^{\infty} dy \; f_{X,Y}(x,y) \tag{3.6}$$

Las funciones de distribución y marginal para la variable Y se obtienen de manera semejante.

El momento n-ésimo de la variable X se define como

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; x^n \; f_{X,Y}(x,y) = \int_{-\infty}^{\infty} x^n \; f_X(x) \; dx$$

Los momentos conjuntos para las variables X e Y se definen como

$$\langle x^n y^m \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; x^n y^m \; f_{X,Y}(x,y) \tag{3.7}$$

El momento conjunto mas comunmente utilizado en física es la **covariancia**

$$\mathrm{Cov}(X,Y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle \tag{3.8}$$

que nos da una medida de la correlación entre fluctuaciones de ambas variables. Altermativamente, suele usarse la **función de correlación** (también llamada coeficiente de Pearson):

$$\mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \, \sigma_Y} \tag{3.9}$$

la cual es adimensional. La función de correlación satisface las siguientes propiedades (facilmente demostrables):

- $\mathrm{Cor}(X,Y) = \mathrm{Cor}(Y,X)$
- $-1 \leq \mathrm{Cor}(X,Y) \leq 1$
- $\mathrm{Cor}(X,X) = 1$, $\mathrm{Cor}(X,-X) = -1$
- $\mathrm{Cor}(aX+b, cY+d) = \mathrm{Cor}(X,Y)$ si $a, c \neq 0$

Dos variables aleatorias X e Y son **independientes** si

$$f_{X,Y}(x,y) = f_X(x) \, f_Y(y) \;\; \forall x, y$$

Si X e Y son independientes se satisfacen las siguientes propiedades

- $\langle x^m y^n \rangle = \langle x^m \rangle \langle y^n \rangle$
- $\mathrm{Cor}(X,Y) = 0$
- $V(X+Y) = V(X) + V(Y)$

Esta última propiedad se generaliza directamente al caso de n variables $X_1, \ldots, X_n$ todas independientes entre sí:

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) \tag{3.10}$$

Al tratar con muchas variables aleatorias, a menudo necesitamos encontrar la densidad de probabilidad de una nueva variable que es función de las anteriores. Por ejemplo, supongamos que conocemos la densidad conjunta $f_{X,Y}(x,y)$ y queremos obtener la densidad de probabilidad de la variable $Z = G(X,Y)$, donde G es una función conocida de dos variables. En una generalización de la Ec.(2.20), tenemos que

$$f_Z(Z) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; \delta\left(z - G(x,y)\right) \; f_{X,Y}(x,y) \tag{3.11}$$

esta expresión se generaliza facilmente al caso de n variables.

Tomemos un ejemplo. Sean X e Y dos variables aleatorias independientes, con distribuciones gaussianas de media nula y variancia unitaria, esto es

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \\ f_Y(y) &= \frac{1}{\sqrt{2\pi}}\, e^{-y^2/2} \end{aligned}$$

Supongamos que queremos conocer la densidad conjunta de las variables

$$\begin{aligned} V &= X + Y \\ W &= X - Y \end{aligned} \tag{3.12}$$

Tenemos que

$$f_{V,W}(v,w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; \delta\left(v - v'(x,y)\right) \; \delta\left(w - w'(x,y)\right) \; e^{-(x^2+y^2)/2}$$

donde $v'(x,y) = x + y$ y $w'(x,y) = x - y$. Cambiando las variables de integración a v', w':

$$f_{V,W}(v,w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dv' \int_{-\infty}^{\infty} dw' \; J\begin{pmatrix} x & y \\ v' & w' \end{pmatrix} \delta\left(v - v'\right) \; \delta\left(w - w'\right) \; e^{-(w'^2+y'^2)/4}$$

donde hemos usado que $x^2 + y^2 = (w'^2 + v'^2)/2$ y

$$J\begin{pmatrix} x & y \\ v' & w' \end{pmatrix} = \frac{1}{2}$$

es el Jacobiano de la transformación (3.12). Así,

$$f_{V,W}(v,w) = \frac{1}{4\pi}\, e^{-(w^2+v^2)/4}$$

Dado que $f_{V,W}(v,w)$ es factorizable, las variables V y W son independientes.

De la Ec.(3.11) la función característica para una variable $Z = G(X,Y)$ resulta

$$\tilde{f}_Z(k) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \; e^{ikG(x,y)} \; f_{X,Y}(x,y) \tag{3.13}$$

3.2. Distribuciones binomiales

3.2.1. Derivación mediante la función característica

Resulta instructivo derivar nuevamente la distribución binomial a partir de los conceptos que hemos visto en esta sección. En un experimento binomial podemos describir el resultado de cada prueba individual mediante una variable aleatoria X_i, donde el índice i corresponde al número de la prueba. La variable X_i puede tomar los valores $x = 1$ con probabilidad p y $x = 0$ con probabilidad $1 - p$. La densidad de probabilidad para esta variable es

$$f_{X_i}(x) = p\,\delta(x-1) + (1-p)\,\delta(x) \tag{3.14}$$

y su función característica resulta

$$f_{X_i}(k) = (1-p) + p\,e^{ik} \tag{3.15}$$

Dado que todas las pruebas son independientes, la densidad de probabilidad conjunta para las variables correspondientes a las n pruebas es

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n)$$

La variable binomial (número de éxitos en las n pruebas) puede expresarse como $Y_n = \sum_{i=1}^{n} X_i$. De la Ec.(3.13), la función característica para la variable Y_n resulta

$$\begin{aligned}
\tilde{f}_{Y_n}(k) &= \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n\; e^{ik(x_1+\cdots+x_n)}\; f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n) \\
&= \tilde{f}_{X_1}(k) \times \cdots \times \tilde{f}_{X_n}(k) = \left(q + p\,e^{ik}\right)^n
\end{aligned} \tag{3.16}$$

donde $q = 1 - p$. Expandiendo el binomio tenemos

$$\tilde{f}_{Y_n}(k) = \sum_{l=0}^{n} \binom{n}{l} p^l q^{n-l} e^{ikl} \tag{3.17}$$

de donde la densidad de probabilidad para la variable Y_n resulta

$$f_{Y_n}(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\; e^{-iky}\; \tilde{f}_{Y_n}(k) = \sum_{l=0}^{n} \binom{n}{l} p^l q^{n-l}\,\delta(y-l) \tag{3.18}$$

que es equivalente a la expressión (2.5).

3.2.2. Límite $n \to \infty$

En muchas aplicaciones vamos a trabajar con distribuciones binomiales con n muy grande. Queremos entonces obtener el comportamiento asintótico de la distribución binomial cuando $n \to \infty$. Tanto el valor medio cono la variancia divergen en este caso. Así, resulta práctico trabajar con la variable aleatoria normalizada

$$Z_n = \frac{Y_n - \langle y_n \rangle}{\sigma_{Y_n}} = \frac{Y_n - pn}{\sqrt{pq\,n}} \tag{3.19}$$

La densidad de probabilidad para la variable Z_n es

$$f_{Z_n}(z) = \int_{-\infty}^{\infty} dy\; \delta\left(z - \frac{y - pn}{\sqrt{pq\,n}}\right)\; f_{Y_n}(y)$$

y la función característica

$$\begin{aligned}
\tilde{f}_{Z_n}(k) &= \int_{-\infty}^{\infty} dz\; e^{ikz}\, f_{Z_n}(z) = \int_{-\infty}^{\infty} dy\; \exp\left(ik\,\frac{y-pn}{\sqrt{pq\,n}}\right)\; f_{Y_n}(y) \\
&= e^{-ik\sqrt{pn/q}}\,\tilde{f}_{Y_n}\left(\frac{k}{\sqrt{pq\,n}}\right) = \left(q\,e^{-ik\sqrt{p/qn}} + p\,e^{ik\sqrt{q/pn}}\right)^n
\end{aligned} \tag{3.20}$$

donde hemos usado la Ec.(3.16) y que $q = 1 - p$. A continuación desarrollamos la cantidad entre paréntesis en potencias de k. Es facil ver que los dos primeros términos no nulos son $1 - k^2/2n$. Así, podemos expresar

$$\tilde{f}_{Z_n}(k) = \left(1 - \frac{k^2}{2n}(1 + R_n)\right)^n \tag{3.21}$$

donde

$$R_n = 2\sum_{m=3}^{\infty} \frac{1}{m!}\left(\frac{ik}{\sqrt{n}}\right)^{m-2} \frac{pq^m + q(-p)^m}{(pq)^{m/2}}$$

Para $n \to \infty$ tenemos que $R_n \to 0$. Sea $Z = \text{lím}_{n\to\infty} Z_n$. Entonces

$$\tilde{f}_Z(k) \equiv \lim_{n\to\infty} \tilde{f}_{Z_n}(k) = \lim_{n\to\infty}\left(1 - \frac{k^2}{2n}\right)^n = e^{-k^2/2} \tag{3.22}$$

Así, en el límite $n \to \infty$ obtenemos

$$f_Z(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\; e^{-ikz}\, e^{-k^2/2} = \frac{1}{\sqrt{2\pi}}\; e^{-z^2/2} \tag{3.23}$$

La variable Z tiene una distribución gaussiana con media nula y variancia $\sigma_Z = 1$. Para el caso $n \gg 1$, pero finito, aún podemos aproximar $f_{Z_n}(z) \approx (1\sqrt{2\pi})\exp(-z^2/2)$. Transformando nuevamente a la variable $y = \sigma_Y\; z + \langle y\rangle$ tenemos finalmente

$$f_{Y_n}(y) \approx \int_{-\infty}^{\infty} dz\; \delta\left(y - \sigma_Y\; z + \langle y\rangle\right) f_Z(z) = \frac{1}{\sigma_Y\sqrt{2\pi}}\; \exp\left(-\frac{(y - \langle y\rangle)^2}{2\sigma_Y^2}\right) \tag{3.24}$$

3.2.3. Caminata aleatoria

El problema de la caminata aleatoria es un ejemplo de un problema binomial en física. Este es un modelo simplificado para el movimiento Browniano, esto es, el movimiento de una partícula bajo la acción de fuerzas aleatorias sin correlación entre sí y constituye la explicación mas básica del fenómeno de difusión. Veremos el caso mas simple de una caminata unidimensional.

Consideremos el caso de una partícula restringida a moverse a lo largo del eje x. A cada intervalo de tiempo τ la partícula tiene una probabilidad $p = 1/2$ de dar un paso de longitud Δ hacia la derecha y una probabilidad $q = 1/2$ de dar un paso equivalente hacia la izquierda. Supongamos que la partícula ha dado N pasos y que los pasos son estadísticamente independientes (la probabilidad de un paso a la derecha o a la izquierda a cada instante es independiente de todos los anteriores). El paso i-ésimo tiene entonces asociada una variable X_i que toma los valores $x = \pm\Delta$ con probabilidades $p = q = 1/2$ y por lo tanto su densidad de probabilidad es

$$f_{X_i}(x) = \frac{1}{2}\,\left(\delta(x - \Delta) + \delta(x + \Delta)\right)$$

y la función característica resulta $\tilde{f}_{X_i}(k) = \cos(k\Delta)$.

El desplazamiento neto Y_N de la partícula al cabo de N pasos esta dado por $Y_N = \sum_{i=1}^{N} X_i$, suponiendo que la partícula parte del orígen. De la Ec.(3.16) tenemos que la función característica para la variable Y_N es

$$\tilde{f}_{Y_N}(k) = \left(\cos(k\Delta)\right)^N \tag{3.25}$$

Vamos a analizar ahora el límite cuando tanto el desplazamiento neto en un paso Δ como el tiempo entre pasos τ se vuelven infinitesimalmente pequeños. En dicho límite tanto Y_N como el

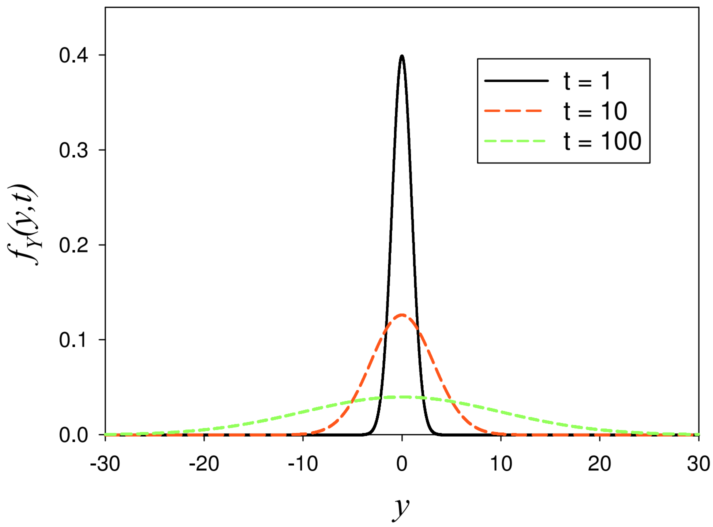

Figura 3.1: Densidad de probabilidad para la caminata aleatoria unidimensional con $D = 1/2$.

tiempo transcurrido $t = N\tau$ pueden aproximarse por variables contínuas. Podemos escribir entonces $\tilde{f}_{Y_N}(k) = \tilde{f}_Y(k,t)$. Dado que la partícula parte del orígen ($y = 0$), tenemos que $f_Y(y, t = 0) = \delta(y)$ y por lo tanto $\tilde{f}_Y(k, 0) = 1$. De la Ec.(3.25) podemos escribir

$$\tilde{f}_Y(k, t+\tau) - \tilde{f}_Y(k,t) = (\cos(k\Delta) - 1)\ \tilde{f}_Y(k,t) = \left(-\frac{k^2\Delta^2}{2} + \cdots\right) \tilde{f}_Y(k,t) \tag{3.26}$$

Si tomamos el límite para $N \to \infty$ y $\tau \to 0$ manteniendo $t = N\tau$ constante, tenemos

$$\lim_{N\to\infty} \lim_{\tau\to 0} \frac{\tilde{f}_Y(k,t+\tau) - \tilde{f}_Y(k,t)}{\tau} = \frac{\partial \tilde{f}_Y(k,t)}{\partial t} \tag{3.27}$$

De la Ec.(3.26) tenemos entonces

$$\frac{\partial \tilde{f}_Y(k,t)}{\partial t} = \lim_{N\to\infty} \lim_{\tau\to 0} \lim_{\Delta\to 0} = \frac{1}{\tau} \left(-\frac{k^2\Delta^2}{2} + \cdots\right) \tilde{f}_Y(k,t) \tag{3.28}$$

La única manera de que este límite sea finito (y distinto de cero) es que

$$\lim_{\tau\to 0} \lim_{\Delta\to 0} \frac{\Delta^2}{\tau} = D < \infty$$

en cuyo caso obtenemos la ecuación diferencial

$$\frac{\partial \tilde{f}_Y(k,t)}{\partial t} = -D\, k^2 \tilde{f}_Y(k,t), \tag{3.29}$$

cuya solución con la condición inicial $\tilde{f}_Y(k,0) = 1$ es

$$\tilde{f}_Y(k,t) = e^{-Dk^2 t} \tag{3.30}$$

La densidad de probabilidad es

$$f_Y(y,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} e^{-Dk^2 t} = \sqrt{\frac{1}{4\pi\, Dt}} \exp\left(-\frac{y^2}{4Dt}\right) \tag{3.31}$$

Así, la probabilidad de encontrar al caminante en una posición entre y e $y + dy$ es gaussiana con media nula y desviación estándar $\sigma = \sqrt{2Dt}$. La densidad de probabilidad se "desparrama"en el tiempo manteniendo el área unitaria, como se muestra en la Fig.3.1. Notemos que $\sqrt{\langle y^2 \rangle} = \sigma \propto \sqrt{t}$. Este es el comportamiento característico de un movimiento difusivo: el desplazamiento de una partícula browniana de su posición inicial crece com $\sqrt{t}$. La constante D se conoce como **coeficiente de difusión**. De hecho, notemos que de la definición

$$f_Y(y,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} \tilde{f}_Y(k,t)$$

tenemos que

$$\frac{\partial f_Y(y,t)}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} \frac{\partial f_Y(k,t)}{\partial t}$$

y

$$\frac{\partial^2 f_Y(y,t)}{\partial y^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-k^2)\, e^{-iky} \tilde{f}_Y(k,t)$$

Así, antitransformando la Ec.(3.29) obtenemos

$$\frac{\partial f_Y(k,t)}{\partial t} = D\, \frac{\partial^2 f_Y(y,t)}{\partial y^2} \tag{3.32}$$

Esta última ecuación diferencial se conoce como **ecuación de difusión** (tiene la misma estructura que la ecuación de difusión de calor) y su solución viene dada por la Ec.(3.31).

3.3. El Teorema del Límite Central

Consideremos un conjunto de N variables aleatórias X_i, todas independientes entre sí e idénticamente distribuidas, esto es, la densidad de probabilidad $f_X(x)$ es la misma para todas. Un ejemplo de esto podría ser la repetición independiente de un mismo experimento N veces. Sean $\mu = \langle x \rangle$ y $\sigma^2 = V(X)$ el valor medio y la variancia correspondientes a la densidad $f_X(x)$.

Consideremos ahora la variable Y_N, definida como la desviación del **promedio aritmético** de las N variables respecto de μ, esto es

$$Y_N \equiv \frac{1}{N} \sum_{i=1}^{N} X_i - \mu = \sum_{i=1}^{N} Z_i \tag{3.33}$$

donde

$$Z_i \equiv \frac{1}{N}(x_i - \mu) \tag{3.34}$$

donde $\langle Z_i \rangle = 0$ y $V(Z_i) = \sigma^2/N^2$. Resulta inmediato que

$$\left\langle \sum_{i=1}^{N} X_i \right\rangle = N\mu$$

con lo cual

$$\langle Y_N \rangle = 0 \tag{3.35}$$

Pensemos en que queremos medir una cierta cantidad, la cual fluctua de experimento en experimento (X_i). Habitualmente queremos estimar el valor medio de esta variable (μ), para lo cual repetimos el experimento N veces y aproximamos μ por el promedio aritmético de las N mediciones. Sin embargo, ese promedio es a su vez una variable aleatoria (si repetimos la serie de N experimentos, el nuevo promedio también fluctua). La Ec.(3.35) nos dice que el valor medio de este promedio coincide con μ, con lo es en principio un estimador aceptable. Sin embargo, querríamos evaluar de alguna manera el error cometido en esta estimación. Para eso tenemos que analizar las fluctuaciones del promedio aritmético respecto de su valor medio, las cuales están descriptas por la variable Y_N. Si recordamos la ecuación (3.10), vemos que, bajo el supuesto de que las mediciones son completamente independientes entre sí, tenemos que

$$V(Y_N) = \sum_{i=1}^{N} V(Z_i) = \frac{\sigma^2}{N} \tag{3.36}$$

con lo cual $\sigma_{Y_N} = \sigma/\sqrt{N}$. Podemos concluir a partir de este resultado que la dispersión del promedio disminuye con N? Estrictamente no, ya que podria ocurrir que momentos de orden superior al segundo se hagan muy grandes. Para dar una respuesta a la pregunta anterior debemos obtener la distribución de probabilidad para la variable Y_N. Para ello calculamos primero la función característica para las variables Z_i. Tenemos que

$$\tilde{f}_Z(k) = \int_{-\infty}^{\infty} dx\, e^{i(k/N)(x-\mu)} f_X(x) = 1 - \frac{1}{2}\frac{k^2}{N^2}\sigma^2 + \cdots \tag{3.37}$$

donde hemos desarrollado en serie de Taylor la exponencial e integrado término a término. Si este desarrollo es válido, es decir, si los momentos de $f_X(x)$ son **finitos**, los término superiores pueden despreciarse para N suficientemente grande (notemos que el caracter oscilatorio del integrando asegura que la integral se va a cero para valores grandes de k). La función característica para Y_N es por lo tanto

$$\tilde{f}_{Y_N}(k) = \left(1 - \frac{1}{2}\frac{k^2}{N^2}\sigma^2 + \cdots\right)^N \to \exp\left(-\frac{k^2\sigma^2}{2N}\right) \quad para \quad N \to \infty \tag{3.38}$$

y la densidad de la variable Y_N:

$$f_{Y_N}(x) \to \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{iky} \exp\left(-\frac{k^2\sigma^2}{2N}\right) = \sqrt{\frac{N}{2\pi\sigma^2}} \exp\left(-\frac{Ny^2}{2\sigma^2}\right) \tag{3.39}$$

Independientemente de la forma particular de $f_X(x)$, si sus momentos son finitos, el promedio de un número grande de mediciones de X tiene una distribución gaussiana centrada en $\langle x \rangle$ y con desvío estándar igual al devío de X dividido por $\sqrt{N}$. Este el el Teorema del Límite Central.