Física Teórica 3 — segundo cuatrimestre de 2023

Guía 3: caminata al azar en tiempo discreto*

■ **Caminata al azar discreta**. Considere el problema de una persona cuya posición puede asumir los valores enteros entre menos y más infinito. Cada segundo, la persona da un paso hacia adelante con probabilidad p, o hacia atrás con probabilidad $q = 1 - p$. Sea $p_m(n)$ la probabilidad de que la persona ocupe la posición m luego de n pasos. La caminata empieza en $t = t_0 = 0$. En general, $t_n = n$.

a) La manera más sencilla de resolver este problema consiste en usar un poco de combinatoria, sumando las probabilidades de caminos disjuntos que llevan, desde el origen, a una misma posición m luego de n pasos: de los n pasos, habrá n_+ hacia adelante y n_- hacia atrás, de modo que $n_+ + n_- = n$. Por otro lado, la posición m será la diferencia $n_+ - n_-$. Notar que n y m determinan unívocamente n_+ y n_-. Lo que resta por averiguar es de cuántas maneras pueden ordenarse los n pasos, dado que n_+ fueron hacia adelante y n_- hacia atrás y cuál es la probabilidad de cada ordenamiento. De este modo, encuentre $p(m, t_n|0, t_0)$ y escriba $p_m(n)$ para cualquier condición inicial $p_m(0)$.

b) La manera complicada de resolver este problema, pero de aplicación más general, empieza por escribir la ecuación de evolución para $p_m(n)$. Esto es: escriba la ecuación que da $p_m(n + 1)$ en términos de $p_m(n)$.

c) A partir de la ecuación de evolución para p_m, escriba y resuelva las ecuaciones de evolución para el valor medio de m y la desviación cuadrática media,

$$\begin{aligned} \overline{m}(n) &= \langle m(n) \rangle\,, \\ \sigma^2(n) &= \left\langle \left[m(n) - \overline{m}(n)\right]^2 \right\rangle = \left\langle m^2(n) \right\rangle - \overline{m}(n)^2 = \overline{m^2}(n) - \overline{m}(n)^2. \end{aligned} \tag{1}$$

d) Calcule $\overline{m}(n)$ y $\sigma^2(n)$ para la caminata simétrica, $p = q = \frac{1}{2}$, cuando la persona parte del origen.

La ecuación de evolución para $p_m(n)$ da una relación de recurrencia en las dos variables, m y n. Para resolverla, puede aplicarse el método de la función generatriz.

e) Defina $F(z, n) = \sum_m p_m(n) z^m$ y escriba su ecuación de evolución. Las probabilidades $p_m(n)$ siempre se leen como los coeficientes que acompañan a z^m en el desarrollo de $F(z, n)$ en potencias de z.

f) Encuentre $F(z, n)$ para una condición inicial arbitraria $p_m(0)$.

g) En el caso particular en el que la persona parte del origen, encuentre $F(z, n)$ y $p_m(n)$.

h) Note que las derivadas respecto de z de la función generatriz, evaluadas en $z = 1$, permiten calcular valores medios. Cuando la persona parte del origen, calculando las derivadas adecuadas, determine $\overline{m}(n)$ y $\sigma^2(n)$ a partir de la función generatriz.

*zanellaj@df.uba.ar

■ **Solución**. Supongamos primero que la persona parte del origen en $t_0 = 0$. Como cada paso es independiente de los otros, la probabilidad de una dada secuencia de pasos es igual al producto de las probabilidades de cada paso. La probabilidad de una caminata con n_+ pasos hacia adelante y n_- hacia atrás es

$$p^{n_+} q^{n_-}. \tag{2}$$

El número de caminatas con estos valores de n_+ y n_- es igual al número de formas en que pueden elegirse los n_+ pasos hacia adelante, o los n_- pasos hacia atrás, dentro del total de $n_+ + n_-$ pasos:

$$N(n_+, n_-) = \binom{n_+ + n_-}{n_+} = \binom{n_+ + n_-}{n_-}. \tag{3}$$

La probabilidad buscada es la suma de las probabilidades de cada caminata, es decir,

$$p(n_+, n_-) = N(n_+, n_-)\, p^{n_+} q^{n_-} = \binom{n_+ + n_-}{n_+} p^{n_+} q^{n_-}. \tag{4}$$

Debemos escribir esto en términos de m y n.

Si la persona está a tiempo t_n en la posición m, debe ser

$$n = n_+ + n_-, \qquad m = n_+ - n_-. \tag{5}$$

Despejando,

$$n_+ = \frac{n+m}{2}, \qquad n_- = \frac{n-m}{2}. \tag{6}$$

Evidentemente, para que estas expresiones tengan sentido, n y m deben tener la misma paridad. Entonces,

$$p(m, t_n|0, t_0) = p\big(\tfrac{n+m}{2}, \tfrac{n-m}{2}\big) = \binom{n}{\frac{n+m}{2}}' p^{\frac{n+m}{2}} q^{\frac{n-m}{2}}. \tag{7}$$

La prima sobre el número binomial indica que si $n + m$ no es par, entonces el número binomial debe tomarse igual a cero. Si inicialmente la distribución es $p_i(0)$, la probabilidad a tiempo t_n de estar en el sitio m es

$$p_m(n) = \sum_{i=-\infty}^{\infty} p(m, t_n|i, t_0)\, p_i(0) = \sum_{i=-\infty}^{\infty} p(m-i, t_n|0, t_0)\, p_i(0). \tag{8}$$

En la última igualdad, usamos la invariancia traslacional: la probabilidad de transición entre los sitios i y j sólo depende de la separación entre los sitios. Explícitamente,

$$p_m(n) = \sum_{i=-\infty}^{\infty} \binom{n}{\frac{n+m-i}{2}}' p^{\frac{n+m-i}{2}} q^{\frac{n-m+i}{2}} p_i(0). \tag{9}$$

Se trata ahora de llevar esta expresión a una forma más simple.

Un número binomial $\binom{k}{l}$ con argumentos enteros y $k \geqslant 0$ es distinto de cero sólo si

$$0 \leqslant l \leqslant k. \tag{10}$$

Esto hace que la Ec. (9) pueda reescribirse como

$$p_m(n) = {\sum_{i=m-n}^{m+n}}' \binom{n}{\frac{n+m-i}{2}} p^{\frac{n+m-i}{2}} q^{\frac{n-m+i}{2}} p_i(0). \tag{11}$$

Ahora la prima sobre el símbolo de la sumatoria indica que i avanza de dos en dos. Con esto ya no es necesario chequear la paridad del argumento inferior del número binomial. Tiene sentido que la probabilidad de estar en m a tiempo n sólo dependa de las probabilidades de estar en el sitio i a tiempo 0 con $m-n \leqslant i \leqslant m+n$, puesto que el sitio m puede alcanzarse en n pasos desde el sitio i sólo si la distancia entre m e i es menor o igual que n. También tiene sentido que la suma avance de dos en dos. Si el sitio m puede alcanzarse desde el sitio i, entonces no puede alcanzarse desde los sitios $i \pm 1$ (demostrarlo).

Volviendo a la Ec. (11), con el cambio de variable

$$j = \frac{n+m-i}{2}, \tag{12}$$

resulta, finalmente,

$$p_m(n) = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} p_{n+m-2j}(0). \tag{13}$$

Este es el resultado principal. Como siempre, es importante verificar que la suma de las probabilidades sea uno. En este caso, tenemos

$$\sum_m p_m(n) = \sum_m \left[\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} p_{n+m-2j}(0) \right] = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} \left[\sum_m p_{n+m-2j}(0) \right]. \tag{14}$$

Debido a que la última suma sobre m se extiende a todos los enteros, podemos hacer el cambio de variable

$$k = n + m - 2j, \tag{15}$$

y la suma sobre k también se extenderá a todos los enteros:

$$\sum_m p_{n+m-2j}(0) = \sum_k p_k(0) = 1. \tag{16}$$

Luego,

$$\sum_m p_m(n) = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} = (p+q)^n = 1. \tag{17}$$

Como aplicación de la Ec. (13), calculemos el valor medio de la posición. Esto es,

$$\begin{aligned}\overline{m}(n) &= \sum_m m p_m(n) = \sum_m m \left[\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j}\, p_{n+m-2j}(0)\right] \\ &= \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} \left[\sum_m m\, p_{n+m-2j}(0)\right].\end{aligned} \tag{18}$$

Ahora bien,

$$\sum_m m\, p_{n+m-2j}(0) = \sum_m (m+2j-n) p_m(0) = \overline{m}(0) + (2j-n). \tag{19}$$

Así,

$$\overline{m}(n) = \left[\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j}\right] \overline{m}(0) + \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} (2j-n). \tag{20}$$

El primer término entre corchetes es igual a $(p+q)^n = 1$, por lo tanto

$$\overline{m}(n) = \overline{m}(0) + \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} (2j-n). \tag{21}$$

Existe un método estándar para calcular este tipo de sumatorias. Notemos que

$$S = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} (2j-n) = \frac{d}{dx}\left[\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} x^{2j-n}\right]_{x=1}. \tag{22}$$

Pero

$$\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} x^{2j-n} = \sum_{j=0}^{n} \binom{n}{j} (px)^j \left(\frac{q}{x}\right)^{n-j} = \left(px + \frac{q}{x}\right)^n. \tag{23}$$

En definitiva,

$$S = \frac{d}{dx}\left[\left(px + \frac{q}{x}\right)^n\right]_{x=1} = n \left(px + \frac{q}{x}\right)^{n-1} \left(p - \frac{q}{x^2}\right)\Big|_{x=1} = n(p-q). \tag{24}$$

De manera que, volviendo a la Ec. (21),

$$\overline{m}(n) = \overline{m}(0) + n(p-q). \tag{25}$$

El valor medio de la distribución se mueve con velocidad $p-q$. Si la caminata es simétrica, $\overline{m}(n)$ es constante. Queda como ejercicio que calculen $\overline{m^2}(n)$ siguiendo el método anterior.

El problema propone otra estrategia para encontrar la probabilidad $p_m(n)$. La idea es escribir la ecuación de evolución para la probabilidad y resolver esta ecuación mediante el método de la función generatriz.

El primer objetivo es escribir una ecuación para $p_m(n+1)$ en términos de $p_m(n)$. Si la persona está a tiempo $n+1$ en la posición m, sólo hay dos maneras en las que pudo llegar allí: i) la persona estaba a tiempo n en la posición $m-1$ y dio un paso hacia adelante, o ii) la persona estaba a tiempo n en la posición $m+1$ y dio un paso hacia atrás. Formalmente,

$$p_m(n+1) = p\, p_{m-1}(n) + q\, p_{m+1}(n). \tag{26}$$

Antes de introducir el método de la función generatriz, vamos a usar esta ecuación para decir algo acerca de los valores medios. Veremos que es posible encontrar $\overline{m}(n)$ y $\sigma^2(n)$ sin necesidad de encontrar explícitamente $p_m(n)$. Por ejemplo, multiplicando la Ec. (26) por m y sumando sobre todo m resulta

$$\sum_m m\, p_m(n+1) = p \sum_m m\, p_{m-1}(n) + q \sum_m m\, p_{m+1}(n). \tag{27}$$

Pero

$$\begin{aligned} \sum_m m\, p_m(n+1) &= \overline{m}(n+1), \\ \sum_m m\, p_{m-1}(n) &= \sum_m (m+1) p_m(n) = \overline{m}(n) + 1, \\ \sum_m m\, p_{m+1}(n) &= \sum_m (m-1) p_m(n) = \overline{m}(n) - 1. \end{aligned} \tag{28}$$

Por lo tanto,

$$\overline{m}(n+1) = (p+q)\overline{m}(n) + p - q = \overline{m}(n) + p - q. \tag{29}$$

Esta ecuación es muy fácil de resolver: a cada paso sumamos $p - q$ al resultado del paso anterior. Empezando desde t_0, obtenemos

$$\overline{m}(n) = \overline{m}(0) + n(p-q). \tag{30}$$

Es el mismo resultado de la Ec. (25).

Por otro lado, para encontrar una ecuación de evolución para la desviación cuadrática media, multiplicamos la Ec. (26) por m^2 y sumamos sobre todo m,

$$\overline{m^2}(n+1) = p \sum_m m^2 p_{m-1}(n) + q \sum_m m^2 p_{m+1}(n). \tag{31}$$

Ahora bien,

$$\begin{aligned} \sum_m m^2 p_{m-1}(n) &= \sum_m (m+1)^2 p_m(n) = \overline{m^2}(n) + 2\overline{m}(n) + 1, \\ \sum_m m^2 p_{m+1}(n) &= \sum_m (m-1)^2 p_m(n) = \overline{m^2}(n) - 2\overline{m}(n) + 1. \end{aligned} \tag{32}$$

Luego,

$$\begin{aligned}\overline{m^2}(n+1) &= p\left[\overline{m^2}(n) + 2\overline{m}(n) + 1\right] + q\left[\overline{m^2}(n) - 2\overline{m}(n) + 1\right] \\ &= \overline{m^2}(n) + 2(p-q)\overline{m}(n) + 1.\end{aligned} \tag{33}$$

La desviación cuadrática media está definida como

$$\sigma^2(n) = \overline{m^2}(n) - \overline{m}(n)^2. \tag{34}$$

Usando las ecuaciones (33) y (29), obtenemos

$$\sigma^2(n+1) = \overline{m^2}(n) + 2(p-q)\overline{m}(n) + 1 - \left[\overline{m}(n) + p - q\right]^2 = \sigma^2(n) + 1 - (p-q)^2. \tag{35}$$

Escribiendo $1 = p + q$, queda

$$1 - (p-q)^2 = (p+q)^2 - (p-q)^2 = 4pq. \tag{36}$$

En definitiva,

$$\sigma^2(n+1) = \sigma^2(n) + 4pq. \tag{37}$$

Esta ecuación también es fácil de resolver,

$$\sigma^2(n) = \sigma^2(0) + 4npq. \tag{38}$$

Si p no es cero ni uno, la desviación cuadrática media es una función creciente. Su velocidad de crecimiento es máxima cuando $p = q = \frac{1}{2}$.

Un resultado clásico es el comportamiento de $\sigma^2(n)$ cuando la caminata es simétrica y la persona parte del origen. En tal caso,

$$\overline{m}(n) = 0, \qquad \sigma^2(0) = 0, \tag{39}$$

y la desviación cuadrática media es

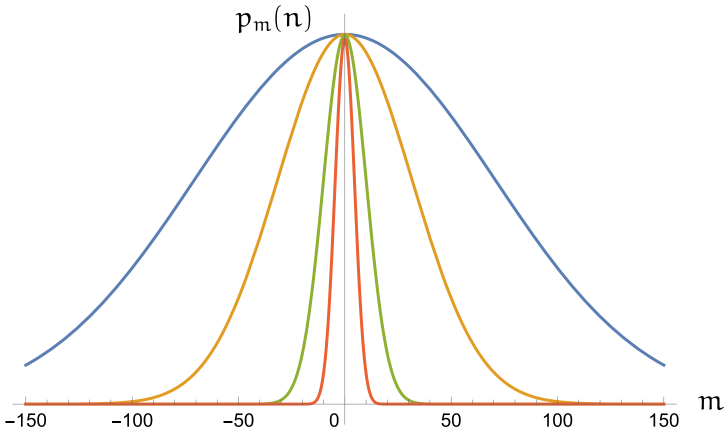
$$\sigma^2(n) = \overline{m^2}(n) = n. \tag{40}$$

La ley de evolución

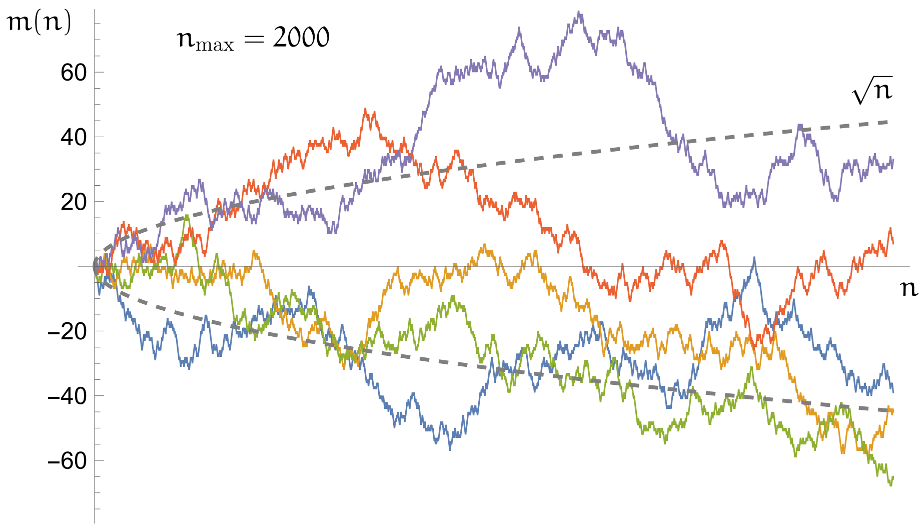
$$\sqrt{\overline{m^2}(n)} = \sqrt{n} \tag{41}$$

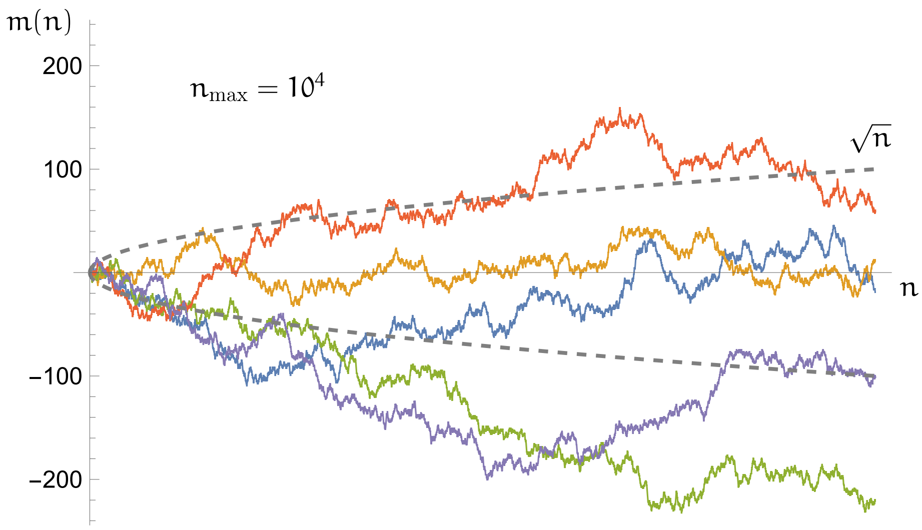
es un resultado que debe recordarse: la distancia media al origen aumenta con el tiempo como $\sqrt{t}$. Esta ley es válida para la caminata al azar en cualquier número de dimensiones y fue uno de los resultados principales del paper de Einstein sobre movimiento browniano de 1905.[†]

[†]Lo pueden bajar aquí. La sección relevante para esta discusión es la 4.



Para la camina simétrica que parte del origen, la figura de arriba muestra la distribución $p_m(n)$ para varios valores de n entre 20 y 5000. El eje vertical está rescaleado para que todas las distribuciones tengan la misma altura. Las figuras de abajo muestran varias realizaciones de la caminata simétrica. La primera figura llega hasta los 2000 pasos; la segunda, hasta los 10 000. Es notable, y contrario a la intuición, el hecho de que sea *frecuente* que las caminatas crucen *rara vez* la línea $m = 0$.





Volvamos al cálculo de $p_n(m)$. La ecuación de evolución

$$p_m(n+1) = p\, p_{m-1}(n) + q\, p_{m+1}(n) \tag{42}$$

es una relación de recurrencia en m y en n. Vamos a definir la función generatriz

$$F(z,n) = \sum_m p_m(n) z^m. \tag{43}$$

La función generatriz contiene toda la información sobre las probabilidades. El motivo para definir esta función es que su ecuación de evolución es fácil de revolver. Multiplicando la Ec. (42) por z^m y sumando sobre todo m, queda

$$F(z,n+1) = p \sum_m p_{m-1}(n) z^m + q \sum_m p_{m+1}(n) z^m. \tag{44}$$

Pero

$$\begin{aligned} \sum_m p_{m-1}(n) z^m &= \sum_m p_m(n) z^{m+1} = zF(z,n), \\ \sum_m p_{m+1}(n) z^m &= \sum_m p_m(n) z^{m-1} = \frac{1}{z} F(z,n). \end{aligned} \tag{45}$$

Así,

$$F(z,n+1) = \left(pz + \frac{q}{z} \right) F(z,n). \tag{46}$$

Con la introducción de la función generatriz, transformamos una relación de recurrencia en dos variables en una relación de recurrencia en una sola variable que tiene una forma especialmente simple. La solución de la Ec. (46) es inmediata,

$$F(z,n) = \left(pz + \frac{q}{z} \right)^n F(z,0). \tag{47}$$

La condición inicial para F está relacionada con la condición inicial para las probabilidades,

$$F(z,0) = \sum_m p_m(0) z^m. \tag{48}$$

Para extraer las probabilidades $p_m(n)$ a partir de la Ec. (47), hay que expandir $F(z,n)$ en potencias de z,

$$F(z,n) = \left[\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} z^{2j-n} \right] \left[\sum_k p_k(0) z^k \right] = \sum_k \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} z^{2j-n+k} p_k(0). \tag{49}$$

Necesitamos aislar los términos que tienen una dada potencia z^m. Es decir, debe ser

$$2j - n + k = m. \tag{50}$$

Esto sugiere hacer el siguiente cambio de variable

$$k = m + n - 2j. \tag{51}$$

La suma sobre m se extiende también a todos los enteros. La Ec. (49) se lee como

$$F(z,n) = \sum_m \left[\sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j}\, p_{m+n-2j}(0) \right] z^m. \tag{52}$$

Luego,

$$p_m(n) = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j}\, p_{n+m-2j}(0), \tag{53}$$

que coincide con la expresión (13).

Si la persona parte del origen en t_0, entonces $p_m(0) = \delta_{m0}$. Por lo tanto, la ecuación anterior se reduce a lo siguiente

$$p_m(n) = \binom{n}{\frac{m+n}{2}}' \, p^{\frac{n+m}{2}} q^{\frac{n-m}{2}}, \tag{54}$$

que es un resultado que ya habíamos encontrado en la Ec. (7).

El método de la función generatriz es muy poderoso. Cuando veamos procesos continuos en el tiempo, también definiremos funciones generatrices, pero esta vez en lugar de satisfacer una relación de recurrencia, satisfarán una ecuación diferencial.

La última parte del ejercicio pide encontrar el valor medio $\overline{m}(n)$ y la desviación cuadrática media $\sigma^2(n)$ a través de la función generatriz. Es fácil ver que la función generatriz tiene las siguientes propiedades:

$$\begin{aligned}
F(1,n) &= \sum_m p_m(n) = 1, \\
\frac{\partial F}{\partial z}(1,n) &= \sum_m m\, p_m(n) = \overline{m}(n), \\
\frac{\partial^2 F}{\partial z^2}(1,n) &= \sum_m m(m-1)\, p_m(n) = \overline{m^2}(n) - \overline{m}(n).
\end{aligned} \tag{55}$$

En el caso de que la persona parta del origen a tiempo t_0, es decir, cuando $p_m(0) = \delta_{m0}$, las Ecs. (47) y (48) dan

$$\begin{aligned}
F(z,0) &= \sum_m \delta_{m0} z^m = 1, \\
F(z,n) &= \left(pz + \frac{q}{z}\right)^n.
\end{aligned} \tag{56}$$

Usando las expresiones (55), obtenemos

$$\overline{m}(n) = n\left(pz + \frac{q}{z}\right)^{n-1}\left(p - \frac{q}{z^2}\right)\Big|_{z=1} = n(p-q),$$

$$\overline{m^2}(n) - \overline{m}(n) = n(n-1)\left(pz + \frac{q}{z}\right)^{n-2}\left(p - \frac{q}{z^2}\right)^2 + 2n\left(pz + \frac{q}{z}\right)^{n-1}\frac{q}{z^3}\Big|_{z=1} \qquad (57)$$

$$= n(n-1)(p-q)^2 + 2nq.$$

Por lo tanto,

$$\overline{m^2}(n) = n(n-1)(p-q)^2 + n. \qquad (58)$$

Entonces,

$$\sigma^2(n) = \overline{m^2}(n) - \overline{m}(n)^2 = n(n-1)(p-q)^2 + n - n^2(p-q)^2$$
$$= n\left[1-(p-q)^2\right] = n\left[(p+q)^2 - (p-q)^2\right] = 4npq. \qquad (59)$$

Esto coincide con el resultado (38) cuando $\sigma^2(0) = 0$. Como última observación, noten que el truco que empleamos en la Ec. (22) para calcular el valor medio de m es una manera encubierta de introducir la función generatriz.