

Capítulo 2

VARIABLES ALEATORIAS Y PROBABILIDAD

Una variable cuyo valor está determinado por la ocurrencia de un evento aleatorio se denomina **variable aleatoria o estocástica**. En otras palabras, una variable aleatoria X es una función del espacio muestral S en los números reales. En un dado experimento, una variable aleatoria puede tomar diferentes valores. Debemos entonces tener cuidado en distinguir entre la variable (que denotaremos con letras mayúsculas) y sus posibles valores $\{x_i\}$ que puede tomar en cada realización del experimento. Por ejemplo, el número de caras que aparece en una tirada de tres monedas es una variable aleatoria X , cuyos posibles valores son $x = 0, 1, 2, 3$.

2.1. Variables aleatorias discretas

Una variable aleatoria X que puede tomar un conjunto numerable (finito o infinito) de valores $X(S) = x_1, x_2, \dots$ se dice **discreta**. Se define la **distribución de probabilidad** $P(x)$ de una variable aleatoria X como la probabilidad de que X tome el valor x , y viene dada por la suma de las probabilidades de todos los puntos muestrales en S para los cuales X toma el valor x . La misma satisface las propiedades

$$0 \leq P(x) \leq 1 \quad \forall x$$

$$\sum_x P(x) = 1$$

Se define el **valor esperado o valor medio** de la variable x como

$$\langle x \rangle \equiv \sum_x x P(x) \tag{2.1}$$

y el mismo representa un promedio pesado de la variable x . De la misma manera, se define el valor medio de una función arbitraria de $g(x)$ como

$$\langle g(x) \rangle \equiv \sum_x g(x) P(x) \tag{2.2}$$

La **variancia** $V(X)$ se define como

$$V(X) \equiv \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \tag{2.3}$$

y el **desvío estándar** como $\sigma_X = \sqrt{V(X)}$. El valor medio nos da una medida del promedio esperable de los valores de X si el experimento se repite muchas veces. El desvío estándar nos da una medida de cuán dispersos estarán estos resultados respecto del valor medio. Desarrollaremos estos conceptos con mayor detalle más adelante.

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

$$\langle x^n \rangle \equiv \sum_x x^n P(x) \quad (2.4)$$

Veremos a continuación algunos ejemplos de distribuciones de probabilidad que aparecen con frecuencia en la práctica.

2.1.1. La distribución de probabilidad binomial

Una de las aplicaciones más comunes de la Teoría de Probabilidades es el caso de un número n muy grande de experimentos, cada uno de los cuales tiene solo dos posibles resultados. Un ejemplo típico es una encuesta de opinión acerca de una votación por un plebiscito (votación por SI o por NO, donde el voto en blanco no está permitido). La empresa encuestadora selecciona una “muestra”, esto es un subconjunto, de n personas dentro de un espacio muestral enorme con $N \gg n$ elementos. Si bien cada persona tiene perfectamente definido su voto, los encuestados son elegidos completamente al azar. Supongamos entonces que una fracción p de los votantes votará por SI. Dado que solo hay dos posibilidades, una fracción $1 - p$ votará por NO. Si se elige entonces una persona al azar, la probabilidad de que vote por SI será justamente p (Ec.(1.3)). La pregunta entonces es: ¿Cuál es la probabilidad de que exactamente x entre los n voten por SI?

Un experimento binomial tiene entonces las siguientes características:

1. El experimento consta de n pruebas idénticas.
2. Cada prueba tiene dos resultados posibles. Llamaremos genéricamente *éxito* E y *fracaso* F .
3. La probabilidad de tener éxito en una sola prueba es p y permanece constante de prueba en prueba (la probabilidad de fracaso por lo tanto es $(1 - p)$).
4. Las pruebas son independientes entre sí.
5. La variable de estudio es X , el número de éxitos observados en las n pruebas.

Otro ejemplo de un experimento binomial sería arrojar n monedas perfectas y contar el número de veces que aparece cara. En este caso $p = 1/2$.

Los puntos muestrales de este experimento consisten en cadenas binarias del tipo

EEEEEEFF...FEFF

Supongamos una cadena particular conteniendo x valores E y $n - x$ valores F y calculemos su probabilidad. Este evento es la intersección de n eventos *independientes*, de los cuales x tienen probabilidad p y $n - x$ tienen probabilidad $1 - p$. Por lo tanto, la probabilidad de la intersección es $p^x(1 - p)^{n-x}$. Ahora bien, esa es la probabilidad de un conjunto *particular* conteniendo x valores E y $n - x$ valores F en un determinado orden. Si cambiamos el orden de las letras, sin alterar la cantidad de letras E y F , obtenemos otro conjunto diferente con la misma probabilidad. Dado que dichos eventos son mutuamente excluyentes, la probabilidad de la unión es la suma de las probabilidades. La probabilidad de obtener x valores E y $n - x$ valores F , sin importar el orden (es lo que buscamos) será entonces $p^x(1 - p)^{n-x}$ multiplicado por el número de *combinaciones* posibles de dichas letras, esto es, C_x^n . Así

$$P(x) = p^x(1 - p)^{n-x} \binom{n}{x} \quad \text{para } x = 0, 1, \dots, n \quad (2.5)$$

Esta es la **distribución binomial**. Podemos verificar que la misma está normalizada ($\sum_x P(x) = 1$):

$$\sum_{x=0}^n P(x) = \sum_{x=0}^n p^x (1-p)^{n-x} \binom{n}{x} = [p + (1-p)]^n = 1$$

donde hemos usado el binomio de Newton Ec.(1.8). Podemos calcular también el valor medio de X :

$$\begin{aligned} \langle x \rangle &= \sum_{x=0}^n x p^x (1-p)^{n-x} \binom{n}{x} \\ &= \sum_{x=1}^n p^x (1-p)^{n-x} \frac{n!}{(n-x)! (x-1)!} \\ &= np \sum_{x=1}^n p^{x-1} (1-p)^{n-x} \frac{(n-1)!}{(n-x)! (x-1)!} \\ &= np \sum_{y=0}^{n-1} p^y (1-p)^{n-1-y} \frac{(n-1)!}{(n-1-y)! y!} \\ &= np \sum_{y=0}^{n-1} p^y (1-p)^{n-1-y} \binom{n-1}{y} \\ &= np \end{aligned} \tag{2.6}$$

En forma semejante se puede demostrar que (Ej. 2.1.4-1)

$$V(X) = np(1-p). \tag{2.7}$$

2.1.2. La distribución de probabilidad geométrica

Supongamos ahora que realizamos una experimento semejante al binomial, pero en lugar de realizar n pruebas, terminamos el experimento cuando aparece por primer vez una E . La variable aleatoria que nos interesa aquí es el número X de la prueba para la cual se obtuvo el primer éxito. La aparición del primer éxito puede tener lugar en la primera prueba, en la segunda o nunca. Así, la variable X en este caso no está acotada. Los elementos del espacio muestral en este caso son: $E_1 = E$ (éxito en la primera prueba), $E_2 = FE$ (éxito en la segunda), $\dots, E_k = FFF \dots E$ (éxito en la k -ésima), etc. Dado que las pruebas son independientes, la probabilidad es

$$P(x) = (1-p)^{x-1} p \quad \text{para } x = 1, 2, \dots, \tag{2.8}$$

Esta se conoce como **distribución geométrica**. Dado que $(1-p) \leq 1$, vemos que la distribución geométrica decae exponencialmente con x , a menos que $p = 1$, en cuyo caso $P(1) = 1$ y $P(x) = 0$ en cualquier otro caso. Verifiquemos la normalización:

$$\sum_{x=1}^{\infty} P(x) = \sum_{x=1}^{\infty} p (1-p)^{x-1} = \frac{p}{1-p} \sum_{x=1}^{\infty} (1-p)^x = \frac{p}{1-p} \left[\frac{1}{[1 - (1-p)]} - 1 \right] = 1$$

donde hemos usado la suma de la serie geométrica $\sum_{x=0}^{\infty} a^x = 1/(1-a)$. Veamos el valor medio ($q \equiv 1-p$):

$$\langle x \rangle = p \sum_{x=1}^{\infty} x q^{x-1} = p \frac{d}{dq} \left(\sum_{x=1}^{\infty} q^x \right) = p \frac{d}{dq} \left(\frac{q}{1-q} \right) = \frac{1}{p}. \tag{2.9}$$

Este resultado es fácil de interpretar. Cuanto menor sea el valor de p , más lentamente decae $P(x)$ y por lo tanto mayor es el valor medio. En forma semejante, puede demostrarse que (Ej. 2.1.4-1)

$$V(X) = \frac{2}{p^2} - \frac{1}{p}. \quad (2.10)$$

2.1.3. La distribución de Poisson

Supongamos que tenemos una sustancia radioactiva y un contador Geiger. La sustancia emite una partícula aleatoriamente en el tiempo cada vez que un átomo decae y marca un conteo en el Geiger. Supongamos que el tiempo de vida media de la sustancia es muy grande comparado con el tiempo de observación, de manera que el número de conteos es relativamente pequeño (no estamos pensando en una reacción en cadena). Cada evento de decaimiento entonces puede considerarse independiente de los otros. Nos interesa entonces determinar la probabilidad de observar X conteos en un intervalo de tiempo dado τ . Podemos llevar a cabo este cálculo subdividiendo el intervalo de tiempo en n subintervalos, de tal manera que τ/n sea suficientemente pequeño para que la probabilidad de que ocurra más de un decaimiento en un subintervalo sea despreciable. Sea p la probabilidad de que ocurra un conteo en un subintervalo. Claramente, p y los resultados subsecuentes van a depender de n . Ahora bien, la manera de independizarnos de la longitud del subintervalo es tomar el límite en que este va a cero. Si bien p depende de la longitud del subintervalo, podemos asumir que es la misma para cualquiera de ellos (para una longitud τ dada). En ese caso, la distribución de probabilidad para X es binomial. Si bien no sabemos en principio como depende p de la longitud del subintervalo, parece razonable que p disminuya a medida que la misma disminuye, es decir, que n aumenta. La dependencia más simple que satisface estos criterios es $p = \lambda/n$, donde λ es una constante. En otras palabras, vamos a asumir que el valor medio de conteos pn se mantiene constante a medida que aumentamos n . La distribución de probabilidad para X se obtiene entonces tomando el límite $n \rightarrow \infty$ de la distribución binomial con $p = \lambda/n$:

$$\begin{aligned} \lim_{n \rightarrow \infty} p^x (1-p)^{n-x} \binom{n}{x} &= \lim_{n \rightarrow \infty} \frac{n(n-1) \cdots (n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \frac{n(n-1) \cdots (n-x+1)}{n^x} \\ &= \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \end{aligned}$$

Todos los factores a la derecha tienden a uno, mientras que

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

Así

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad (2.11)$$

Esta se conoce como **distribución de Poisson**. Esta distribución se aplica en general cada vez que tenemos un proceso que ocurre aleatoriamente en un intervalo de tiempo o espacio, cuando la probabilidad de ocurrencia de un evento es constante e independiente de los demás eventos. Algunos otros ejemplos de aplicación son el número de autos que pasan a través de un cierto punto en una ruta durante un periodo definido de tiempo, el número de llamadas telefónicas en una central telefónica por minuto, el número de animales muertos encontrados por unidad de longitud de ruta, etc. La distribución fue descubierta por Siméon-Denis Poisson (1781-1840) quien la publicó, junto

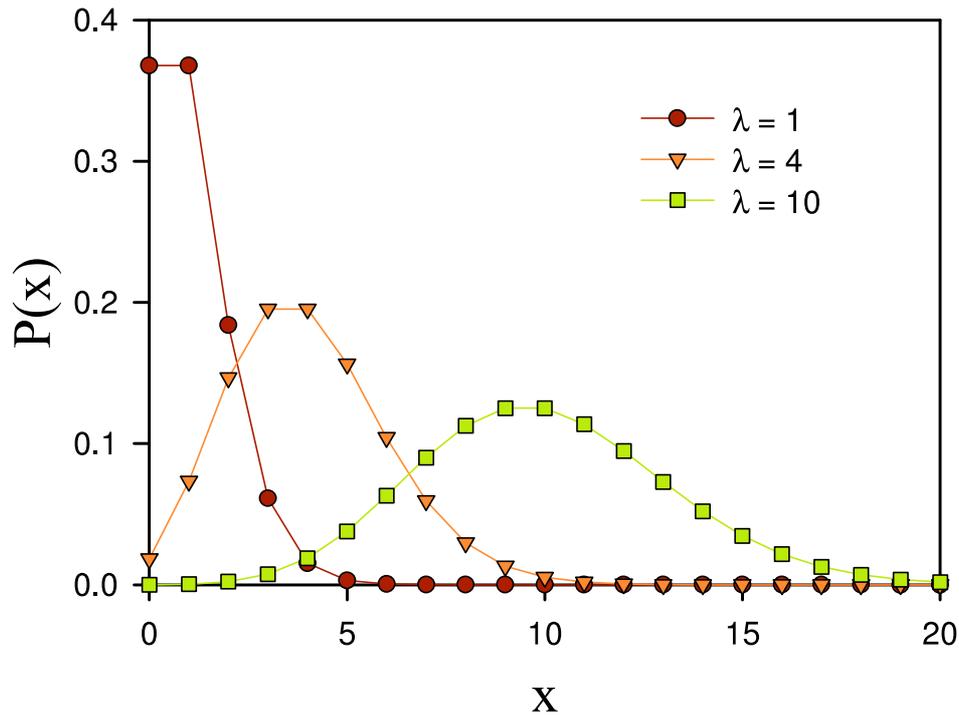


Figura 2.1: Distribución de Poisson para diferentes valores del valor medio λ .

con su teoría de probabilidad, en 1838 en su trabajo *Recherches sur la probabilité des jugements en matières criminelles et matière civile* ("Investigación sobre la probabilidad de los juicios en materias criminales y civiles").

Verifiquemos la normalización:

$$\sum_{x=0}^{\infty} P(x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda} e^{-\lambda} = 1$$

y el valor medio

$$\langle x \rangle = e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda \quad (2.12)$$

De la misma manera se puede demostrar que (Ej. 2.1.4-1)

$$V(X) = \lambda. \quad (2.13)$$

En la Fig.2.11 se muestran algunos ejemplos de la distribución de Poisson para diferentes valores de λ .

2.1.4. Ejercicios

1. Demuestre las ecuaciones (2.7), (2.10) y (2.13)

2.2. Variables aleatorias continuas

Una variable aleatoria X que puede tomar un conjunto no-numerable de valores en un dado intervalo del eje real se dice continua. Un intervalo (a, b) contenido en el dominio de definición de la

variable corresponde entonces a un evento. Vamos a introducir entonces una **densidad de probabilidad** $f_X(x)$, tal que la probabilidad de que X tome valores dentro del intervalo comprendido entre x y $x + dx$ esta dada por $f_X(x) dx$. La probabilidad de que X tome valores en un intervalo finito (a, b) viene entonces dada por

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx$$

La densidad de probabilidad debe ser continua a tramos, satisfacer $f_X(x) \geq 0$ y

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

Notemos que, de acuerdo con esta definición, la probabilidad de que una variable continua tome un valor bien definido $P(X = x) = 0$.

Podemos englobar también dentro de esta definición a las variables aleatorias discretas. Si X es una variable aleatoria discreta que toma valores x_1, x_2, \dots , con probabilidad $p_i = P(x_i)$, entonces su densidad de probabilidad será

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

donde $\delta(x)$ es la función delta de Dirac.

Se define la **Función de distribución** $F_X(x)$ (no hay que confundirla con la distribución de probabilidad de una variable discreta) como la probabilidad $P(X \leq x)$:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(x') dx'$$

Se sigue entonces que $f_X(x) = dF_X(x)/dx$. Dado que f_X es no-negativa, la función de distribución es siempre no-decreciente. Por la normalización de f_X , la función de distribución toma los valores límite $F_X(-\infty) = 0$ y $F_X(\infty) = 1$. Para el caso de una variable aleatoria discreta tendremos que

$$F_X(x) = \sum_i p_i \Theta(x - x_i)$$

donde $\Theta(x)$ es la función escalón de Heaviside, esto es, $\Theta(x) = 0$ para $x \leq 0$ y $\Theta(x) = 1$ para $x > 0$.

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

$$\langle x^n \rangle \equiv \int_{-\infty}^{\infty} x^n f_X(x) dx$$

Veamos un poco la interpretación de estas cantidades. Todas las propiedades de la variable X están contenidas en la función densidad (algo así como una “relación fundamental” para la variable). La probabilidad de que X tome valores en un dado intervalo está dada por el área bajo f_X para ese intervalo.

El primer momento $\langle x \rangle$ (**media o valor medio**) nos dá el “centro de masa” de la densidad f_X . Esta cantidad a menudo se confunde con otras dos cantidades: el **valor mas probable** x_p y la **mediana** x_m . El valor mas probable de X se define como el máximo de f_X . La mediana se define como el valor de x que divide el área bajo la curva $f_X(X)$ en partes iguales. En otras palabras, $F_X(x_m) = 1/2$. En algunos casos estas cantidades coinciden (como veremos mas adelante) pero en general son diferentes.

El segundo momento $\langle x^2 \rangle$ nos dá el “momento de inercia” de la densidad f_X respecto del origen. El desvío estándar $\sigma_X = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$ nos dá una medida de cuán lejos se dispersa la probabilidad respecto de la media $\langle x \rangle$, esto es, el desvío medio cuadrático.

Supongamos una variable tal que $\langle x \rangle = 0$ (siempre podemos hacer que esto ocurra desplazando el origen, esto es, restando la media de la variable). El tercer momento $\langle x^3 \rangle$ nos dá una medida de cuán asimétrica es la distribución respecto del origen. Esto es, si la densidad es simétrica respecto del origen, es una función par y por lo tanto el tercer momento se anula. Cuanto mayor sea el tercer momento, menos “simétrica” será f_X .

Veamos algunos ejemplos. Un variable tiene distribución **uniforme** si

$$f_X(x) = \begin{cases} A & \text{si } a \leq x \leq b \\ 0 & \text{en otro caso} \end{cases} \quad (2.14)$$

donde A es una constante. La normalización exige que $A = 1/(b - a)$. La probabilidad de que X tome valores en un intervalo (c, d) , con $a < c < d < b$ resulta $(d - c)/(b - a)$. Un cálculo directo muestra que $\langle x \rangle = (b + a)/2$, esto es, el centro del intervalo. De la misma forma es fácil mostrar que $\sigma_X = (b - a)/\sqrt{12}$. Un cálculo directo nos muestra que la probabilidad de que

$$P(\langle x \rangle - \sigma_X \leq x \leq \langle x \rangle + \sigma_X) \approx 0,58$$

Esto es, el intervalo $\pm \sigma_X$ alrededor de la media concentra aproximadamente el 60% de la probabilidad. La función de distribución se obtiene también fácilmente:

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ A(x - a) & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases} \quad (2.15)$$

De aquí puede verificarse inmediatamente que $x_m = \langle x \rangle$.

2.2.1. Distribución de Gauss

Otro ejemplo de gran importancia es la **distribución de Gauss** o **normal**, definida por:

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (2.16)$$

definida para todo x real, con $\sigma > 0$. El lector puede verificar por integración directa que la misma está normalizada y que $\langle x \rangle = \mu$, $\sigma_X = \sigma$. Esta curva tiene su máximo en $x = \mu$ y es simétrica respecto del valor medio. Por lo tanto en este caso la media, la mediana y el valor más probable coinciden. Mediante integración numérica (o mediante valores de tablas), puede verificarse que la probabilidad de que la variable tome valores en un intervalo $\pm \sigma$ alrededor de la media es aproximadamente 68%, mientras que para un intervalo $\pm 2\sigma$ la probabilidad es de aproximadamente 95%.

2.2.2. Distribución exponencial

Un ejemplo de una distribución no simétrica es la **distribución exponencial**

$$f_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{\mu} e^{-x/\mu} & \text{si } x \geq 0 \end{cases} \quad (2.17)$$

El lector puede verificar por integración directa que la misma está normalizada y que $\langle x \rangle = \mu$, $\sigma_X = \mu$. Esta curva tiene su máximo en $x = 0$. Por otra parte, la función de distribución es

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x/\mu} & \text{si } x \geq 0 \end{cases} \quad (2.18)$$

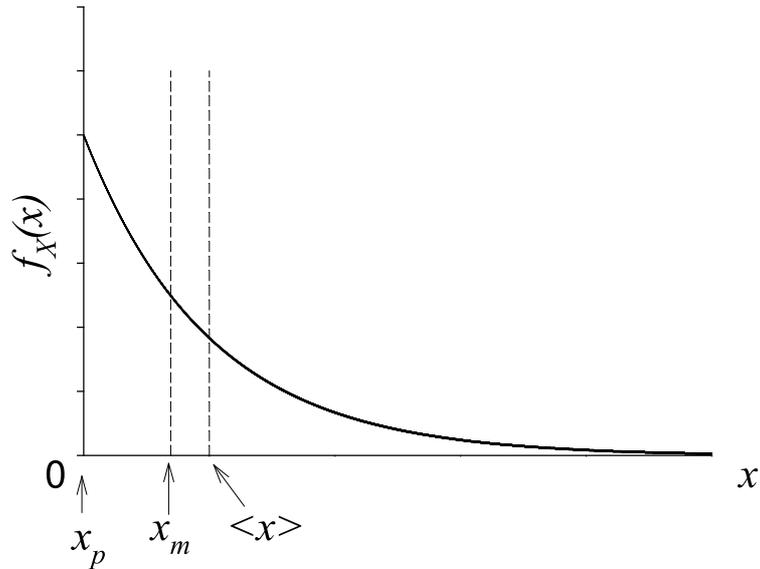


Figura 2.2: Distribución exponencial, media, mediana y valor mas probable.

Si resolvemos la ecuación $F(x_m) = 1/2$, obtenemos $x_m = \mu \ln 2$. Vemos que en este caso la media, la mediana y el valor mas probable no coinciden. Los mismos se muestran en la Fig.2.2 junto con la densidad.

2.2.3. Transformación de variables aleatorias

Sea $Y = g(X)$, donde $g(x)$ es una función arbitraria y X una variable aleatoria. Siendo X aleatoria, Y también lo es. Supongamos que conocemos la densidad de probabilidad $f_X(x)$. Nos preguntamos entonces cual es la densidad de probabilidad $f_Y(y)$. Para relacionar ambas cantidades notemos que

$$\langle y \rangle = \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_{-\infty}^{\infty} y f_Y(y) dy \quad (2.19)$$

Resulta inmediato verificar que

$$f_Y(y) = \int_{-\infty}^{\infty} \delta(y - g(x)) f_X(x) dx \quad (2.20)$$

satisface la condición anterior. Usando la propiedad de la delta de Dirac:

$$\delta(h(z)) = \sum_i \frac{1}{|h'(z_i)|} \delta(z - z_i)$$

si $h'(z_i) \neq 0$, donde z_i son los ceros de $h(z)$, resulta

$$f_Y(y) = \sum_i \left| \frac{dg_i^{-1}(y)}{dy} \right| f_X[g_i^{-1}(y)] \quad (2.21)$$

donde $g_i^{-1}(y)$ son las diferentes ramas de la función inversa de $g(x)$ (estamos suponiendo el caso general en que la función $g(x)$ no tiene una inversa única). La Ec.(2.21) puede interpretarse fácilmente. Supongamos primero que la función $g(x)$ tiene inversa única. Entonces

$$f_Y(y) = \left| \frac{dx(y)}{dy} \right| f_X[x(y)]$$

Si $g(x)$ es creciente ($g'(x) > 0$) y por lo tanto $dx(y)/dy > 0$ entonces

$$f_Y(y) dy = f_X(x) dx$$

Esta ecuación nos dice que la probabilidad de que Y tome valores entre y e $y + dy$ es igual a la probabilidad de que X tome valores entre x y $x + dx$, donde $dy = g'(x) dx$. Si $g(x)$ es decreciente, su derivada es negativa y por lo tanto

$$f_Y(y) dy = -f_X(x) dx$$

ya que $f_Y(y)$ debe ser positiva. Si la función $g^{-1}(y)$ es multivaluada, entonces podemos repetir las consideraciones anteriores, pero para evaluar la probabilidad de que Y tome valores entre y e $y + dy$, debemos sumar las probabilidades de todas las ramas de $g^{-1}(y)$ correspondientes al intervalo de X entre x y $x + dx$.

2.2.4. Función característica y desarrollo en cumulantes

La función característica $\tilde{f}_X(k)$ correspondiente a una variable aleatoria X se define como

$$\tilde{f}_X(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} e^{ikx} f_X(x) dx \quad (2.22)$$

esto es, la transformada de Fourier de $f_X(x)$. Desarrollando en serie de Taylor la exponencial e integrando término a término tenemos:

$$\tilde{f}_X(k) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle x^n \rangle}{n!} \quad (2.23)$$

Es importante notar que el desarrollo anterior solo es válido si los momentos decrecen suficientemente rápido con n como para que la serie converja; en otras palabras, el intercambio entre la serie y la integración no siempre es válido. En el caso en que este desarrollo sea válido, vemos que si tenemos todos los momentos podemos reconstruir la densidad de probabilidad $f_X(x)$ antitransformando

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} \tilde{f}_X(k) dk \quad (2.24)$$

$\tilde{f}_X(k)$ es una función continua de k , compleja con las propiedades $\tilde{f}_X(k=0) = 1$, $|\tilde{f}_X(k)| \leq 1$ y $\tilde{f}_X^*(k) = \tilde{f}_X(-k)$ (* denota complejo conjugado). Si conocemos la función característica, podemos obtener los momentos por diferenciación:

$$\langle x^n \rangle = \lim_{k \rightarrow 0} (-i)^n \frac{d^n \tilde{f}_X(k)}{dk^n}$$

Tomemos por ejemplo la distribución de Gauss Ec.(2.16). De la definición (2.22) completando cuadrados en el exponente es fácil ver que

$$\tilde{f}_X(k) = e^{ik\mu - k^2\sigma^2/2} \quad (2.25)$$

Derivando una vez obtenemos:

$$\langle x \rangle = \lim_{k \rightarrow 0} (-i) (i\mu - k\sigma^2) \tilde{f}_X(k) = \mu$$

Derivando dos veces obtenemos:

$$\langle x^2 \rangle = \lim_{k \rightarrow 0} (-i)^2 [-\sigma^2 + (i\mu - k\sigma^2)^2] \tilde{f}_X(k) = \sigma^2 + \mu^2$$

Alternativamente al desarrollo en potencias de la función característica, podemos desarrollar en potencias el *logaritmo* de la misma. esto es

$$\ln \tilde{f}_X(k) = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \quad (2.26)$$

Esto se conoce como **desarrollo en cumulantes**, donde el coeficiente de la potencia n -ésima $C_n(X)$ se conoce como *cumulante de orden n* . De la definición, tenemos que

$$\tilde{f}_X(k) = \exp \left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \right] \quad (2.27)$$

Usando el desarrollo en serie de Taylor de la exponencial en la Ec.(2.27) y el desarrollo (2.23) e igualando potencias de k , podemos expresar los cumulantes en términos de los momentos. Así, por ejemplo:

$$\begin{aligned} C_1(X) &= \langle x \rangle \\ C_2(X) &= \langle x^2 \rangle - \langle x \rangle^2 = V(X) \\ C_3(X) &= \langle x^3 \rangle - 3 \langle x \rangle \langle x^2 \rangle + 2 \langle x \rangle^3 \end{aligned}$$

y en general puede verse que el *cumulante de orden n* es función de todos los momentos de orden $l \leq n$. Si comparamos la función característica (2.25) para la distribución de Gauss con la expresión (2.27) vemos que para la misma se anulan todos los cumulantes de orden $n > 2$. Esto significa que todos los momentos de orden superior a 2 pueden ser expresados en función de los dos primeros momentos. Esto es una particularidad de la distribución de Gauss.