CAPÍTULO 10
VALORES ESPERADOS, FUNCIONES GENERADORAS DE MOMENTOS Y MOMENTOS

10.1 Introducción

Después de haber estudiado las bases de la teoría de la probabilidad, las variables aleatorias, las funciones de probabilidad y las funciones de densidad, en este capítulo estudiaremos la otra forma de caracterizar a dichas distribuciones a partir de sus parámetros mediante los cuales es posible representar la localización de la distribución, la dispersión de la variable aleatoria, el sesgo y el aplanamiento; estos parámetros que pueden encontrarse si se conocen las funciones generadoras de los momentos. Iniciaremos este capítulo con un concepto central de la teoría de la probabilidad conocido como la esperanza matemática o el valor esperado de la variable aleatoria. Conviene indicar desde ahora que distinguiremos dos tipos de parámetros; a los primeros las llamaremos los parámetros particulares que están comprendidos dentro de las funciones de probabilidad o sea que forman parte de ella lo que hace ver que son familias de distribuciones y cada conjunto de valores, que puede ser uno o varios, seleccionan un miembro de la familia.

Por otro lado, los segundos parámetros, son los parámetros generales caracterizan las medidas de localización, dispersión, sesgo y aplanamiento; deben calcularse a partir de los momentos y contienen a los parámetros particulares, como se verá en el desarrollo de los temas de este capítulo.

10.2 Valores Esperados o Esperanzas de Variables Aleatorias y de Funciones de Variables Aleatorias

En general, el valor esperado es una suma pesada de los valores de la variable aleatoria o de una función de variables aleatorias, por las probabilidades de dichas variables o de las funciones –ver figura 10.1-; que se denota por $E[X]$, significa que $E[X]$ ES UN OPERADOR que actúa sobre la VA X o sobre la función de la VA y, dependiendo del tipo de variable aleatoria, se define como:

$E[X] = \sum_{i=1}^{\infty}(x_i)p(x_i)$, para las VA's discretas; y (10.1)

$E[X] = \int_{-\infty}^{+\infty} xf(x)dx$, para las VA's continuas. (10.2)

Desde el punto de vista matemático es necesario tener presente que: a) la sumatoria o la integral puede no converger a un valor finito, en cuyo caso, el valor esperado no existe; sin embargo, para los propósitos de este libro, siempre existirá la convergencia y b) el *valor esperado es un operador que actúa sobre variables aleatorias o funciones de variables aleatorias*.

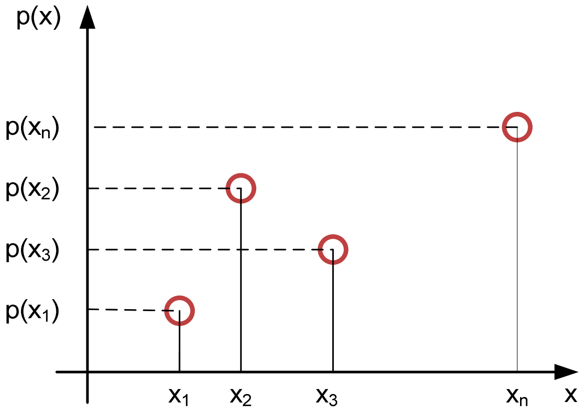


Figura 10.1 Distribución de probabilidad de una VA discreta

Una primera interpretación del valor esperado se ilustra con los siguientes ejemplos.

Ejemplo 10.1 Supóngase que el Director de PEMEX Exploración Producción debe decidir si en determinada región, debe o no perforar un pozo. Los eventos posibles sin A = {sí existe petróleo} y B = {no existe petróleo}, sin importar el volumen. Si toma la decisión de perforar y encuentra petróleo el valor del petróleo obtenido será de \$250, el costo de la perforación costará \$50 y el beneficio neto será entonces de \$200. Por el contrario, si decide no perforar, puede someter a concurso con las compañías petroleras la perforación con las siguientes bases: \$25 por los derechos de perforación más otros \$25 si se descubre petróleo. El Director tiene la incertidumbre sobre el hallazgo del petróleo por lo que se asesora del grupo de geólogos de PEP que son más expertos que él en la exploración quienes, basados en la información existente y las pruebas geológicas que practican, asignan la probabilidad de 0.25 de que sí se encontrará petróleo en dicha región (todos los precios son figurados y están en millones de pesos).

Con esta información se obtiene la siguiente tabla de pagos, comúnmente utilizada en la teoría de decisiones:

Tabla X. matriz de pagos

		Estado de la	**naturaleza (X)**
Decisión por tomar		I: SÍ existe petróleo	II: NO existe petróleo
	Sí perforar	200	-50
	No perforar	50	25

Para evaluar esta situación, se aplica el valor esperado para las dos alternativas de decisión:

Si decide perforar: $E[X/\mathrm{I}] = 200(0.25) - 50(0.75) = 12.50$
Si decide no perforar: $E[X/\mathrm{II}] = 50(0.25) + 25(0.75) = 31.25$

Por lo tanto, a la luz de estos resultados, la decisión debe ser No perforar. Cabe observar que una cosa es el valor esperado y otra es el valor deseado.

Ejemplo 10.2 La proporción de impurezas de una muestra de petróleo crudo se modela por la función de densidad:

$$f(x) = f(x) = \begin{cases} cx^2 + x, & 0 \leq x \leq 1 \\ 0 & otros\ casos \end{cases}$$

Para poder determinar el valor esperado de X, necesitamos primeramente encontrar el valor de c.
Sabemos que $\int_{-\infty}^{\infty} f(x)dx = 1$
Con lo cual, para nuestro caso:

$$\int_0^1 (cx^2 + x)dx = [c\frac{x^3}{3} + \frac{x^2}{2}]\Big|_0^1 = \left(c\frac{1}{3} + \frac{1}{2}\right) - 0 = 1, \text{ entonces } c = \frac{3}{2}.$$

Sustituyendo el valor de c en la función de densidad, calculemos el valor esperado:

$$E[X] = \int_0^1 x(\frac{3}{2}x^2 + x)dx = \int_0^1 (\frac{3}{2}x^3 + x^2)dx = \left[\frac{3}{2} \times \frac{x^4}{4} + \frac{x^3}{3}\right]_0^1 = \frac{3}{8} + \frac{1}{3} = \frac{17}{24} = 0.71$$

Que corresponde a la proporción de impurezas esperadas, que en la estadística matemática significa el valor promedio de cualquier variable aleatoria sobre un número indefinido de muestreos, como se verá más adelante.

10.2.1 El álgebra de los valores esperados

Como se observa en (10.1) y (10.2), el valor esperado E es una suma ponderada de los valores de X por sus probabilidades correspondientes, por lo que las reglas algebraicas de los valores esperados son extensiones de las reglas de las sumatorias -las integrales son sumas- que se aplican a las VA' discretas y continuas. Por la utilidad que tienen en la probabilidad, es altamente recomendable que se familiaricen con ellas.

Regla 1
Si c es una constante real, entonces, el valor esperado de la constante es la constante misma:

$$E[c] = c \tag{10.3}$$

Demostración:

Usando (10.2), $E[c] = \int_{-\infty}^{+\infty} cf(x)dx = c\int_{-\infty}^{+\infty} f(x)dx = c \times 1 = c$, puesto que $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Ejemplo 10.3 Si se ganará \$10 independientemente del evento que ocurra, la ganancia esperada es \$10.

Regla 2

Si c es una constante real y el valor esperado de X es $E[X]$, entonces:

$$E[cX] = cE[X] \tag{10.4}$$

Demostración:

Usando (10.2), $E[cX] = c\int_{-\infty}^{+\infty} xf(x)dx = cE[X]$

Ejemplo 10.4 Si el valor esperado del tiempo de arribo de los aviones al aeropuerto de la Ciudad de México es $E[T]$ y se sabe que una persona arribará el quinto aterrizaje, entonces el valor esperado del tiempo de espera es $5E[T]$.

Regla 3

Si c es una constante real y X es una variable aleatoria, entonces:

$$E[X + c] = E[X] + c \tag{10.5}$$

Demostración:

Usando (2) y la Regla 1:

$$E[X + c] = \int_{-\infty}^{+\infty} (x + c)f(x)dx = \int_{-\infty}^{+\infty} xf(x)dx + c\int_{-\infty}^{+\infty} f(x)dx = E[X] + C$$

Ejemplo 10.5 Si el costo de la producción de un artículo depende del costo de los artículos fabricados (X) y de un costo fijo, representado por $C = X + 2$, se tiene $E[C] = E[X + 2] = E[X] + 2$; es decir, el valor esperado del costo depende del valor esperado de los artículos fabricados más el costo fijo de los indirectos que para nuestro caso es 2.

Las siguientes reglas devienen de uso del valor esperado como operador que se aplica sobre su argumento una prueba más formal implica la transformación de variables.

Regla 4

Si X e Y son variables aleatorias y sus valores esperados son $E[X]$ y $E[Y]$, respectivamente, entonces:

$$E[X + Y] = E[X] + E[Y] \tag{10.6}$$

Ejemplo 6. Si un caballero invita a su novia a comer en un restaurante, y si las funciones de distribución del precio de los platillos son:

Costo de los platillos	Para el caballero: X	Para su novia: Y
120	0.2	0.1
150	0.5	0.35
200	0.3	0.55

El valor esperado de lo que tendrá que pagar el caballero es

$E[X+Y]=E[X]+E[Y]$
$E[X]=120\times 0.2+150\times 0.5+200\times 0.3=159$
$E[Y]=120\times 0.1+150\times 0.35+200\times 0.55=174.5$, con lo cual:
$E[X+Y]=159+174.5=333.5$

Si, además paga 10% de propina, entonces:

$E[1.1X+1.1Y]=1.1\times E[X+Y]=1.1(159+174.5)=1.1\times 333.5=366.85$

Cabe observar dos cosas: a) que $E[X]\neq E[Y]$ porque las distribuciones de X e Y son diferentes y b) la regla 4 se mantiene para cualquier relación entre X e Y.

Ejemplo 10.7 La función de costos del ejemplo de la Regla 3 fuera
$C=Y=g(X)=5e^{X+2}$, entonces

$E[X+Y]=E[X+5e^{X+2}]=E[X]+E[5e^{X+2}]=E[X]+5E[e^{X+2}]$

Debe tenerse cuidado al aplicar el operador valor esperado E a su argumento, pues en el ejemplo anterior $5E[e^{X+2}]\neq 5e^{E[X+2]}$, igual sucede con funciones tales como $E[(X-5)^2]\neq (E[X]-E[5])^2$ y $E[\sqrt[3]{X+c}\,]\neq \sqrt[3]{E[X]+c}$.

<u>Regla 5</u>
Si $g(X)$ es una función de X, entonces, para variables aleatorias discretas

$$E[g(X)]=\sum_{i=1}^{\infty} g(X)P(x_1) \tag{10.7}$$

y para las continuas,

$$E[g(X)]=\int_{-\infty}^{\infty} g(x)f(x)dx \tag{10.8}$$

Ejemplo 10.8 Si la función de costos del ejemplo anterior es $g(X)=C=X+2$, entonces:

$E[g(X)]=E[x+2]=E[X]+E[2]=E[X]+2$

<u>Regla 6</u>

Dado un número finito de variables aleatorias, el valor esperado de la suma de las variables aleatorias es igual a la suma de sus valores esperados, o sea,

$$E[X_1 + X_2 + X_3 + \cdots + X_n] = E[X_1] + E[X_2] + E[X_3] + \cdots + E[X_n] \tag{10.9}$$

En general, para cualquier combinación lineal de variables aleatorias

$$Y = a_1X_1 + a_2X_2 + X_3a_3 + \cdots + a_nX_n,$$

donde las a_i son constantes, se tiene

$$\begin{aligned} E[Y] &= E[a_1X_1 + a_2X_2 + a_3X_3 + \cdots + a_nX_n] \\ &= a_1E[X_1] + a_2E[X_2] + a_3E[X_3] + \cdots + a_nE[X_n] \end{aligned} \tag{10.10}$$

Ejemplo 10.9 Si $Y = 4X^4, Y = -3X^3 y\ Z = 2e^{-5X}$ se tiene

$$E[X + Y + Z] = E[4X^4 - 3X^3 + 2e^{-5X}] = 4E[X^4] - 3E[X^3] + 2E[e^{-5X}].$$

Obsérvese la aplicación de E como operador matemático.

10.3 Funciones Generadoras de Momentos

La metodología para generar momentos es una herramienta potente de la teoría de la probabilidad, porque es menos complicada que el cálculo directo a partir de las funciones de probabilidad.

Existen muchos métodos para generar los momentos que son útiles para diversas tipos de variables aleatorias; así, la Función Generadora de Momentos Factoriales es útil para las variables aleatorias no negativas y enteras; a su vez, la Función Generadora de Momentos se usa para las variables aleatorias discretas o continuas y negativas o no negativas, para las cuales se puede definir esta función; en tanto que la Función Característica está definida para todas las variables aleatorias. Todas estas Funciones son útiles para encontrar la distribución y los momentos de sumas y promedios pesados de variables aleatorias independientes; finalmente, con la Transformada de Mellin es posible determinar la distribución y los momentos de productos o razones de variables aleatorias no negativas o independientes. Por otro lado. Aquí presentaremos la Función Generadora de Momentos Factoriales y la Función Generadora de Momentos por sus bastas aplicaciones.

10.3.1 La Función Generadora de Momentos Factoriales (FGMF) - $G_X(\theta)$-

Como ya se anticipó, si la variable aleatoria (VA) X es discreta cuyos valores posibles son enteros y no negativos, como sucede con las distribuciones binomial, binomial negativa, geométrica, hipergeométrica, Poisson y otras -todas las cuáles se estudiarán en otro capítulo posterior-; entonces, para cualquier número real θ para el cual $0 \leq \theta \leq 1$, se define la Función Generadora de Momentos Factoriales de la distribución de la VA X como:

$$G_X(\theta) = E(\theta^X) = \sum_{k=0}^{\infty} \theta^k p_X(k) \tag{10.11}$$

Como $0 \leq \theta \leq 1$, $0 \leq \theta^k \leq 1$ y $0 \leq k < \infty$; con lo cual $0 \leq G_X(\theta) \leq 1$ es una función bien definida. Cabe observar el valor esperado E opera sobre θ^X. A los momentos generados por esta función se les representan con $\boldsymbol{\mu_{(r)}}$.

Ejemplo 10.10 Calculemos la FGMF de la distribución Binomial.

Si $p_X(k/n,p) = \binom{n}{k} p^k (1-p)^{n-k}$ es la distribución Binomial con parámetros n y p, al sustituirla en (3):

$$G_X(\theta) = E(\theta^X) = \sum_{k=0}^{\infty} \theta^k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{\infty} \binom{n}{k} (\theta p)^k (1-p)^{n-k}$$

$$G_X(\theta) = [\theta p + (1-p)]^n \tag{10.11}$$

Ejemplo 10.11 Calculemos la FGMF de la distribución de la distribución de Poisson $p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$. Al sustituirla en (10.3) tenemos:

$$G_X(\theta) = E(\theta^X) = \sum_{k=0}^{\infty} \theta^k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} \frac{(\theta\lambda)^k}{k!} e^{(-\lambda+\theta\lambda-\theta\lambda)}$$

$$G_X(\theta) = e^{-\lambda+\theta\lambda} \sum_{k=0}^{\infty} \frac{(\theta\lambda)^k}{k!} e^{-\theta\lambda} = e^{-\lambda+\theta\lambda} \tag{10.12}$$

10.3.1.1 Propiedades

Propiedad 1: Unicidad

La Función Generadora de Momentos Factoriales determina unívocamente la función masa de probabilidad de cualquier VA definida sobre los números enteros no negativos:

$$G_X(\theta) = p_X(0) + \sum_{k=1}^{\infty} \theta^k p_X(k)$$

Si $\theta = 0$, $G_X(0) = p_X(0)$ y puede demostrarse que:

$$\mu_{(k)} = p_x(k) = \frac{1}{k!} G_X^{(k)}(0), para\ k = 1,2,3 \ldots;$$

Donde $G_X^{(k)}(0) = \frac{d^k}{d\theta} G_X(0)$; es decir, $\mu_{(k)} = p_X(k)$ se encuentra calculando la derivada de orden k de $G_X(\theta)$ respecto a θ evaluándola para $\theta = 0$ y multiplicándola por $\frac{1}{k!}$.

Ejemplo 10.12 Para la distribución binomial (4) encontremos su FMP.

Vimos –de (10.4)- que $G_X(\theta) = [\theta p + (1-p)]^n$

Calculando la primera derivada de $G_X(\theta)$ y avaluándola en $\theta = 0$ se tiene:

$G_X^{(1)}(0) = \frac{d^1}{d\theta} G_X(0) = \frac{d^1}{d\theta}[\theta p + (1-p)]^n|_{\theta=0} = n[\theta p + (1-p)]^{n-1} p|_{\theta=0} =$

$np(1-p)^{n-1}$

Con lo cual:

$\mu_{(1)} = p_x(1) = \frac{1}{1!} G_X^{(1)}(0) = \frac{1}{1!} np(1-p)^{n-1} = np(1-p)^{n-1}.$

$G_X^{(2)}(0) = \frac{d^2}{d\theta} G_X(0) = \frac{d^2}{d\theta}[\theta p + (1-p)]^n|_{\theta=0} = n(n-1)[\theta p + (1-$
$pn-2p2\theta=0$

$G_X^{(2)}(0) = n(n-1)p^2(1-p)^{n-2}$, con lo cual:

$\mu_{(2)} = p_x(2) = \frac{1}{2!} G_X^{(2)}(0) = \frac{1}{2} n(n-1)p^2(1-p)^{n-2}.$

Siguiendo el mismo procedimiento:

$\mu_{(3)} = p_x(3) = \frac{1}{3!} G_X^{(3)}(0) = \frac{1}{6} n(n-1)(n-2)p^3(1-p)^{n-3}$

En general:

$$\mu_{(k)} = p_x(k) = \frac{1}{k!} G_X^{(k)}(0) = \frac{1}{k!} \frac{n!}{(n-k)!} p^k (1-p)^{n-k} \tag{10.13}$$

Propiedad 2: Los momentos factoriales

Si X es una VA entera no negativa con FMP $p_X(k)$ cuya Función Generadora de Momentos Factoriales es $G_X(\theta)$, y si el r momento factorial de X

$$E[X(X-1)(X-2)\dots(X-r+1)] = \sum_{k=0}^{\infty} k(k-1)(k-2)\dots(k-r+1)p_X(k)$$

$$E[X(X-1)(X-2)\dots(X-r+1)] = \sum_{k=r}^{\infty} k(k-1)(k-2)\dots(k-r+1)p_X(k)$$

es finito, entonces:

$$G_X^{(r)}(\mathbf{1}) = E[X(X-1)(X-2)\dots(X-r+1)]$$

Ejemplo 10.13 Para la distribución de Poisson se obtuvo $G_X(\theta) = e^{-\lambda+\theta\lambda}$ -ecuación (10.5)- se tiene:

$$G_X^{(1)}(\theta) = \frac{d^1}{d\theta} G_X(\theta) = \frac{d^1}{d\theta} e^{-\lambda+\theta\lambda} = \lambda e^{-\lambda+\theta\lambda}$$

$$G_X^{(2)}(\theta) = \frac{d^2}{d\theta} G_X(\theta) = \frac{d^2}{d\theta} e^{-\lambda+\theta\lambda} = \lambda^2 e^{-\lambda+\theta\lambda}$$

$$G_X^{(3)}(\theta) = \frac{d^3}{d\theta} G_X(\theta) = \frac{d^3}{d\theta} e^{-\lambda+\theta\lambda} = \lambda^3 e^{-\lambda+\theta\lambda}$$

Y, en general:

$$G_X^{(r)}(\theta) = \frac{d^r}{d\theta} G_X(\theta) = \frac{d^r}{d\theta} e^{-\lambda+\theta\lambda} = \lambda^r e^{-\lambda+\theta\lambda}$$

Entonces $G_X^{(r)}(1) = E[X(X-1)(X-2)\dots(X-r+1)] = \lambda^r$ y

$$\mu_{(1)} = G_X^{(1)}(1) = E[X] = \lambda$$

$$\mu_{(2)} = G_X^{(2)}(1) = E[X(X-1)] = \lambda^2$$

$$\mu_{(3)} = G_X^{(3)}(1) = E[X(X-1)(X-2)] = \lambda^3; \text{ y}$$

$$E[X(X-1)] + E[X] - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Propiedad 3: Combinaciones lineales
Si a y b son constantes enteras positivas, entonces:

$$G_{aX+b} = E[\theta^{aX+b}] = \theta^b E[(\theta^a)^X] = \theta^b G_X(\theta^a) \tag{10.14}$$

Ejemplo 10.14 Si X se distribuye conforme una distribución geométrica con parámetro p, su FGMF es $G_X(\theta) = \frac{\theta p}{[1-(1-P)\theta]}$; ahora bien, SI $Y = X - 1$ tiene una distribución binomial negativa con parámetros $u = 1$ y p, se tiene

$$G_X(\theta) = G_{Y+1}(\theta) = \theta G_Y(\theta)$$

Y la FGMF de la distribución Binomial Negativa con dichos parámetros es:

$$G_Y(\theta) = \frac{1}{\theta} G_X(\theta) = \frac{1}{\theta} \frac{\theta p}{[1-(1-P\theta)]} = \frac{p}{1-(1-P)\theta} \tag{10.15}$$

Propiedad 4: Convolución
Dadas las VA's X e Y independientes, enteras y no negativas con Funciones Generadoras de Momentos Factoriales $G_X(\theta)$ y $G_Y(\theta)$, respectivamente, entonces se satisface:

$$G_{X+Y}(\theta) = G_X(\theta) G_Y(\theta) \tag{10.16}$$

Ejemplo 10.15 Si X_1, X_2, $X_3, \dots, X_n$ son VA's mutuamente independientes que se distribuyen conforme a una distribución de Poisson con parámetros λ_i, $i = 1,2,3..n$; y

FGMF's $G_{X_1}(\theta), G_{X_2}(\theta), G_{X_3}(\theta) \dots G_{X_n}(\theta)$, respectivamente; la suma de las VA's es $Y = X_1 + X_2 + X_3 + \dots + X_n$; aplicando la propiedad de la Convolución se tiene:

$$G_Y(\theta) = G_{X_1}(\theta)G_{X_2}(\theta)G_{X_3}(\theta) \dots G_{Xn}(\theta)$$

$$G_Y(\theta) = e^{-\lambda_1+\theta\lambda_1}e^{-\lambda_2+\theta\lambda_2}e^{-\lambda_3+\theta\lambda_3} \dots e^{-\lambda_n+\theta\lambda_n} = e^{(-\sum_1^n \lambda_i+\theta\sum_1^n \lambda_i)}$$

La VA Y se distribuye conforme la distribución de Poisson con parámetro $\lambda = \sum_{i=1}^{n} \lambda_i$

Propiedad 5: Preservación de los límites

Para cada $n = 1,2,3, \dots$ sea $p_x(k/n)$ la función masa de probabilidad de los enteros no negativos de k y $G_X(\theta/n)$ su correspondiente FGMF; entonces, si

$$\lim_{n\to\infty} G_X(\theta/n) = G_X(\theta) \tag{10.17}$$

Donde $G_X(\theta)$ es la FGMF correspondiente a la FMP $p_X(k)$, se cumple que:

$$\lim_{n\to\infty} p_X(k/n) = p_X(k),\ \ k = 0, 1,2, \dots$$

Ejemplo 10.16 Ya vimos arriba que para la distribución Binomial $G_X(\theta) = [\theta p + (1-p)]^n$-ecuación (10.4)- y, para la de Poisson $G_X(\theta) = e^{-\lambda+\theta\lambda}$ -ecuación (10.5)-; aplicando esta propiedad de preservación de los límites demostremos que, cuando n es muy grande, entonces la probabilidad Binomial $p(k|n, p_n)$ con parámetros n y $p_n = \frac{\lambda}{n}$ se aproxima a las probabilidades de la distribución de Poisson $p(k) = \frac{\lambda^k e^{-k}}{k!}$ con parámetro λ.

Para la Binomial, $G_X(\theta/n) = [\theta p_n + (1-p_n)]^n = [1 + (\theta-1)p_n]^n$

Cuando $|(\theta-1)p_n| \leq 1$,

$$logG_X(\theta/n) = nlog[1 + (\theta-1)p_n]$$

$$= n(\theta-1)p_n - \frac{n(\theta-1)^2 {p_n}^2}{2} + \frac{n(\theta-1)^3 {p_n}^3}{3} - \cdots$$

Como $\lim_{n\to\infty} np_n = \lambda, \lim_{n\to\infty} np_n^k = 0\ para\ k > 1\ entonces$:

$\lim\limits_{n\to\infty} logG_X(\theta/n) = (\theta-1)\lambda$, de donde:

$$\lim_{n\to\infty} G_X\left(\frac{\theta}{n}\right) = e^{\left(\lim\limits_{n\to\infty} logG_X\left(\frac{\theta}{n}\right)\right)} = e^{(\theta-1)\lambda} = e^{\lambda\theta-\lambda} = G_X(\theta)$$

y, como consecuencia de esta propiedad:

$$\lim_{n\to\infty} p(k/n, p_n) = p(k) == \frac{\lambda^k e^{-k}}{k!}$$

Lo que muestra que las probabilidades Binomiales $p(k/n, p_n)$ para n grandes se aproximan a las de Poisson $p(k)$. Cabe observar que esta prueba no usa el hecho que $p_n = \frac{\lambda}{n}$, sino solamente la consideración $\lim_{n\to\infty} np_n = \lambda$.

10.3.2 Función Generadora de Momentos Ordinarios (FGM) -$M_X(\theta)$-

A diferencia de la FGMF, esta función genera los momentos ordinarios de cualquier variable aleatoria X, ya sea discreta o continua, que son de suma utilidad para calcular los parámetros característicos de las distribuciones; y se define como sigue:

$$M_X(\theta) = E\left[e^{\theta X}\right] \tag{10.18}$$

Para todos los valores de θ en los que existe la FGM.
Aplicando el operador valor esperado E:

Si X es una VA discreta: $M_X(\theta) = E\left[e^{\theta X}\right] = \sum_{i=1}^{\infty}\left(e^{\theta X}\right)p(x_i)$ (10.19)

Si X es una VA continua: $M_X(\theta) = E\left[e^{\theta X}\right] = \int_{-\infty}^{+\infty} e^{\theta X} f(x)dx$ (10.20)

Los momentos ordinarios generados por esta FGM se llaman momentos respecto al origen y se representan con μ'_k

Obsérvese que para $\theta = 0$:

$$\mu'_0 = M_X(0) = E[e^{0X}] = \sum_{i=1}^{\infty} p(x_i) = \int_{-\infty}^{+\infty} f(x)dx = E[1] = 1$$

Y, cuando $M_X(\theta)$ está definida en el intervalo abierto (θ_I, θ_S) que contiene $\theta = 0$, la FGM tiene las propiedades que describen la distribución de X, similares a las descritas para la FGMF $G_X(\theta)$; cuando X es una VA entera no negativa, entonces M_X existe para $-\infty < \theta \leq 0$, y

$$M_X(\theta) = E\left[e^{\theta X}\right] = E\left[(e^{\theta})^X\right] = G_X\left(e^{\theta}\right) \tag{10.21}$$

La principal ventaja de la FGM sobre la FGMF es el significado que tiene sobre una amplia gama de variables aleatorias, en tanto que su desventaja consiste en que no siempre existe en el intervalo abierto (θ_I, θ_S) de valores que contiene a θ; sin embargo, esto no sucede en las distribuciones de probabilidad que analizaremos en este texto.

Ejemplo 10.17 Calculemos las FGM de las distribuciones Binomial, de Poisson y de la Geométrica.

Como se vio en el Ejemplo 10.12, la FGMF de la distribución Binomial es (10.4):

$G_X(\theta) = [\theta p + (1-p)]^n$; Entonces, aplicando (10.9) se tiene:

$$M_X(\theta) = G_X(e^\theta) = [e^\theta p + (1-p)]^n \tag{10.22}$$

De igual forma, en el Ejemplo 10.12 se demostró que, para la distribución de Poisson:

$G_X(\theta) = e^{-\lambda+\theta\lambda}$ y su FGM es - aplicando (10.9)-

$$M_X(\theta) = G_X(e^\theta) = e^{-\lambda+e^\theta\lambda} \tag{10.23}$$

Y para la distribución Geométrica $G_X(\theta) = \frac{\theta p}{[1-(1-\theta P)]}$, con la cual - aplicando (10.9) nuevamente-:

$$M_X(\theta) = G_X(e^\theta) = \frac{e^\theta p}{[1-(1-e^\theta P)]} \tag{10.24}$$

Ejemplo 10.18 Si Z es una variable aleatoria continua con función de densidad de probabilidad $f(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt[2]{2\pi}}$; entonces, su función generadora de momentos es:

$$M_Z(\theta) = E[e^{\theta Z}] = \int_{-\infty}^{+\infty} e^{\theta Z}\frac{e^{-\frac{z^2}{2}}}{\sqrt[2]{2\pi}}dx = \int_{-\infty}^{+\infty} e^{\theta Z}\frac{e^{-\frac{z^2}{2}}}{\sqrt[2]{2\pi}}dx$$

$$= \int_{-\infty}^{+\infty}\frac{e^{-\frac{z^2}{2}+\theta Z}}{\sqrt[2]{2\pi}}dx = e^{\frac{\theta^2}{2}}\int_{-\infty}^{+\infty}\frac{e^{-\frac{1}{2}(z-\theta)^2}}{\sqrt[2]{2\pi}}dx = e^{\frac{\theta^2}{2}}$$

$$M_Z(\theta) = e^{\frac{\theta^2}{2}} \tag{10.25}$$

Al igual que para la FGMF, la FGM también tiene propiedades similares que se enunciarán a continuación.

10.3.2.1 Propiedades de la FGM

Propiedad 1: Unicidad

A $M_X(\theta)$ le corresponde una y solo una distribución; es decir, la FGM caracteriza a una y solo una distribución. Puesto que la Función de Distribución Acumulativa $F_X(x)$ determina la distribución de cualquier VA, a $M_X(\theta)$ le corresponde una y solo una $F_X(x)$.

Ejemplo 10.19 A la distribución Binomial $p_X(k/n,p)=\binom{n}{k}p^k(1-p)^{n-k}$, le corresponde únicamente la FGM: $M_X(\theta)=G_X(e^\theta)=[e^\theta p+(1-p)]^n$. Y para la distribución $f(z)=\frac{e^{-\frac{z^2}{2}}}{\sqrt[2]{2\pi}}$, su única FGM es $M_Z(\theta)=e^{\frac{\theta^2}{2}}$.

Propiedad 2: Momentos

$$\mu'_r=M_X^{(r)}(\mathbf{0})=E[X^r] \tag{10.26}$$

Donde $M_X^{(r)}(\mathbf{0})=\frac{d^r}{d\theta}M_X(\mathbf{0})$.

En efecto, $M_X^{(r)}(\theta)=(\frac{d}{d\theta})^r M_X(\theta)=(\frac{d}{d\theta})^r E\left[e^{\theta X}\right]=E\left[\left(\frac{d}{d\theta}\right)^r e^{\theta X}\right]=E\left[X^r e^{\theta X}\right]$

Para $\theta=0$

$$\mu'_0=M_X^{(r)}(\mathbf{0})=E\left[X^r e^{\theta X}\right]=E[X^r]$$

Propiedad 2' Los Momentos respecto al origen

Además de las funciones de distribución de las VA's, los momentos de una distribución de probabilidades son extremadamente útiles para encontrar los parámetros que la caracterizan conocidos como medidas de localización, dispersión, sesgo y aplanamiento. Los momentos se definen en términos de los valores esperados de las diferentes potencias de las distancias al origen de los valores de la variable aleatoria X, en cuyo caso se les llaman los momentos respecto al origen –ver la Tabla 1-; o bien, con respecto a otro punto c diferente del origen, conocidos como momentos respecto a c *o momentos centrales.* La siguiente tabla muestra los momentos respecto al origen de las VA's discretas y continuas.

Cabe observar que la dificultad en el cálculo de los momentos estriba en la complejidad de las funciones de densidad $f(x)$, que ciertamente no son sencillas o lo complicado de los cálculos de las sumatorias, por lo cual muchas veces se prefiere utilizar las funciones generadoras de momentos de las distribuciones que las caracterizan; es decir, para cada Función Masa de Probabilidad FMP o Función de Densidad de Probabilidad FDP le corresponde una y solo una Función Generadora de Momentos. Antes de definir dichas funciones presentamos algunos ejemplos.

Tabla 10.1 Momentos respecto al origen

Momento	Notación	Para VA's discretas	Para VA's continuas
primer	μ'_1	$\sum_{i=1}^{n} x_i p(x_i)$	$\int_{-\infty}^{\infty} x f(x)dx$
segundo	μ'_2	$\sum_{i=1}^{n} {x_i}^2 p(x_i)$	$\int_{-\infty}^{\infty} x^2 f(x)dx$

tercer	μ_3'	$\sum_{i=1}^{n} x_i{}^3 p(x_i)$	$\int_{-\infty}^{\infty} x^3 f(x)dx$
...	...	...	...
De orden r	μ_n'	$\sum_{i=1}^{n} x_i{}^r p(x_i)$	$\int_{-\infty}^{\infty} x^r f(x)dx$

Ejemplo 10.20 Para el ejercicio del caballero que invita a comer a su novia, se tienen los segundos momentos:

Para la novia: $E[Y^2] = \mu'_{2/Y} = 120^2 \times 0.1 + 150^2 \times 0.35 + 200^2 \times 0.55 = 26{,}130$

Para el novio: $E[X^2] = \mu'_{2/X} = 120^2 \times 0.2 + 150^2 \times 0.5 + 200^2 \times 0.3 = 31{,}315$

Ejemplo 10.21 Para el ejemplo de la proporción de impurezas en una muestra de petróleo crudo, el tercer momento es:

$$\mu_3' = E[X^3] = \int_0^1 x^3(\frac{3}{2}x^2 + x)dx = \int_0^1 (\frac{3}{2}x^5 + x^4)dx = \left[\frac{3}{2}\times\frac{x^6}{6} + \frac{x^5}{5}\right]\begin{matrix}1\\0\end{matrix} = \frac{1}{4} + \frac{1}{5} = \frac{9}{20}$$

Ejemplo 10.22 En el Ejemplo 10.20 demostramos que si Z es una variable aleatoria continua con función de densidad de probabilidad $f(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt[2]{2\pi}}$ -que es la función de densidad normal estándar-; su función generadora de momentos es $M_Z(\theta) = e^{\frac{\theta^2}{2}}$, calculemos sus cuatro primeros momentos:

$$\mu_1' = M_X^{(1)}(\mathbf{0}) = \frac{d^1}{d\theta}M_X(\mathbf{0}) = |_{\theta=0} = \theta e^{\frac{\theta^2}{2}}|_{\theta=0} = 0$$

$$\mu_2' = M_X^{(2)}(\mathbf{0}) = \frac{d^2}{d\theta}M_X(\mathbf{0}) = \frac{d^2}{d\theta}e^{\frac{\theta^2}{2}}|_{\theta=0} = \frac{d^1}{d\theta}\theta e^{\frac{\theta^2}{2}}|_{\theta=0} = (\theta^2 e^{\frac{\theta^2}{2}} + e^{\frac{\theta^2}{2}})|_{\theta=0} = 1$$

$$\mu_3' = M_X^{(3)}(\mathbf{0}) = \frac{d^3}{d\theta}e^{\frac{\theta^2}{2}}|_{\theta=0} = \frac{d}{d\theta}(\theta^2 e^{\frac{\theta^2}{2}} + e^{\frac{\theta^2}{2}})|_{\theta=0} = (\theta^3 e^{\frac{\theta^2}{2}} + 2\theta e^{\frac{\theta^2}{2}} + \theta e^{\frac{\theta^2}{2}})|_{\theta=0} = 0$$

$$\mu_4' = M_X^{(4)}(\mathbf{0}) = \frac{d^4}{d\theta}e^{\frac{\theta^2}{2}}|_{\theta=0} = \frac{d}{d\theta}(\theta^3 e^{\frac{\theta^2}{2}} + 2\theta e^{\frac{\theta^2}{2}} + \theta e^{\frac{\theta^2}{2}})|_{\theta=0}$$

$$\mu_4' = M_X^{(4)}(\mathbf{0}) = \left[e^{\frac{\theta^2}{2}}(3 + 3\theta) + e^{\frac{\theta^2}{2}}(3\theta^2 + \theta^4)\right]|_{\theta=0} = e^{\frac{\theta^2}{2}}(3 + 6\theta + \theta^4)|_{\theta=0} = 3$$

Propiedad 3: Transformación Lineal

Si $Y = cX + d$ es una transformación lineal de X,

$$M_Y(\theta) = M_{cX+d}(\theta) = e^{d\theta}E[c\theta] \qquad (10.27)$$

En efecto, $M_Y(\theta) = M_{cX+d}(\theta) = E\left[e^{(cX+d)\theta}\right] = e^{d\theta}E\left[e^{(c\theta)X}\right] = e^{d\theta}E[c\theta]$

Propiedad 4: Convolución

Para las VA's Xe Ycon FGM's $M_X(\theta) y\, M_Y(\theta)$ se tiene:

$$M_{X+Y}(\theta) = M_X(\theta)M_Y(\theta) \tag{10.28}$$

En efecto, si Xe Yson continuas e independientes:

$M_{X+Y}(\theta) = E\left[e^{\theta(X+Y)}\right] = E\left[e^{\theta X}\right]E\left[e^{\theta Y}\right] = M_X(\theta)M_Y(\theta)$; o bien:

$$M_{X+Y}(\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\theta(X+Y)} f_X(x)f_Y(y)dxdy = \int_{-\infty}^{\infty} e^{\theta X} f_X(x)dx \int_{-\infty}^{\infty} e^{\theta Y} f_Y(y)dy$$

$$M_{X+Y}(\theta) = M_X(\theta)M_Y(\theta)$$

Y, en general, para $X_1, X_2, X_3, \dots, X_n$ variables aleatorias mutuamente independientes y $a_1, a_2, a_3, \dots, a_n$constantes, $Y = \sum_{i=1}^{n} a_i X_i$ se tiene que

$$M_Y(\theta) = M_{X_1}(a_1\theta)M_{X_2}(a_2\theta)M_{X_3}(a_3\theta) \dots M_{X_n}(a_n\theta)$$

Ejemplo 10.23 Sean X y Z VA's que se distribuyen normalmente con medias y varianzas $X \sim N(\mu, \sigma^2)$ y $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$; como vimos en el ejemplo 10.20 -expresión (10.19)- $M_Z(\theta) = e^{\frac{\theta^2}{2}}$ y, además, $X = \sigma Z + \mu$; entonces, por la propiedad 3:

$$M_X(\theta) = M_{\sigma Z+\mu}(\theta) = e^{\mu\theta} M_Z(\theta\sigma) = \left(e^{\mu\theta}\right)\left(e^{\frac{(\theta\sigma)^2}{2}}\right)$$

$$M_X(\theta) = e^{\mu\theta + \frac{\theta^2\sigma^2}{2}}$$

Ejemplo 10.24 Calculemos los cuatro primeros momentos de la variable aleatoria X que se distribuye conforme la distribución Normal $X \sim N(\mu, \sigma^2)$, a partir de fu FGM.

$M_X^{(1)}(\theta) = \frac{d\left(e^{\mu\theta + \frac{\theta^2\sigma^2}{2}}\right)}{dx} = \left(e^{\mu\theta + \frac{\theta^2\sigma^2}{2}}\right)(\mu + \theta\sigma^2)$, avaluando en $\theta = 0$,

$\mu_1' = M_X^{(1)}(0) = \mu;$

$$M_X^{(2)}(\theta) = \frac{d^2\left(e^{\mu\theta + \frac{\theta^2\sigma^2}{2}}\right)}{dx} = \frac{\frac{d}{dx}\left(e^{\mu\theta + \frac{\theta^2\sigma^2}{2}}\right)(\mu + \theta\sigma^2)}{dx}$$

$M_X^{(2)}(\theta) = \left(e^{\mu\theta + \frac{\theta^2\sigma^2}{2}}\right)[\sigma^2 + (\mu + \theta\sigma^2)^2]$; evaluando en $\theta = 0$,

$$\mu_2' = M_X^{(2)}(0) = \sigma^2 + \mu^2$$

De igual forma,

$$M_X^{(3)}(\theta) = \frac{d(e^{\mu\theta+\frac{\theta^2\sigma^2}{2}})[\sigma^2+(\mu+\theta\sigma^2)^2]}{dx} = (e^{\mu\theta+\frac{\theta^2\sigma^2}{2}})[3\sigma^2(\mu+\theta\sigma^2)+(\mu+\theta\sigma^2)^3]; \text{ y}$$

$$\mu_3' = M_X^{(3)}(0) = 3\sigma^2 + \mu^3$$

Finalmente:

$$M_X^{(4)}(\theta) = \frac{d(e^{\mu\theta+\frac{\theta^2\sigma^2}{2}})[3\sigma^2(\mu+\theta\sigma^2)+(\mu+\theta\sigma^2)^3]}{dx} =$$

$$= (e^{\mu\theta+\frac{\theta^2\sigma^2}{2}})[3\sigma^4 + 6\sigma^2(\mu+\theta\sigma^2)^2 + (\mu+\theta\sigma^2)^4]; \text{ evaluando en } \theta = 0,$$

$$\mu_4' = M_X^{(4)}(0) = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4$$

Ejemplo 10.25 Sean X e Y dos VA's que se distribuyen conforme a la distribución Normal, con medias y varianzas $X{\sim}N(\mu_X, \sigma_X^2)$ y $Y{\sim}N(\mu_X, \sigma_X^2)$; por la propiedad de convolución:

$$M_{X+Y}(\theta) = M_X(\theta)M_Y(\theta) = (e^{\mu_X\theta+\frac{\theta^2\sigma_X^2}{2}})(e^{\mu_Y\theta+\frac{\theta^2\sigma_Y^2}{2}}) = e^{[(\mu_X+\mu_Y)\theta+\frac{1}{2}(\sigma_X^2+\sigma_Y^2)\theta^2}$$

Por la propiedad 1 se tiene,

$$X + Y{\sim}N(\mu_X + \mu_Y,\ \sigma_X^2 + \sigma_Y^2)$$

La propiedad 4 puede generalizarse para $X_1, X_2, X_3, \dots, X_n$ VA's mutuamente independientes y $a_1, a_2, a_3, \dots, a_n$ constantes, para la combinación lineal $Y_n = \sum_{i=1}^{n} a_i X_i$ se tiene:

$$M_Y(\theta) = M_{Y_n}(\theta) = M_{X_1}(a_1\theta)M_{X_2}(a_2\theta)M_{X_3}(a_3\theta) \dots M_{X_n}(a_n\theta) \qquad (10.29)$$

Ejemplo 10.26. Si las VA's $X_1, X_2, X_3, \dots, X_n$ son mutuamente independientes con la misma distribución de probabilidades, entonces, por la propiedad de unicidad tienen la misma FGM $M_X(\theta)$; y la FGM de la media $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, por la propiedad 4, es:

$$M_{\bar{X}}(\theta) = [M_X\left(\frac{\theta}{n}\right)]^n$$

Y la Función generadora de Momentos para la población total $S_n = \sum_{i=1}^{n} X_i$ es:

$M_{S_n}(\theta) = [M_X(\theta)]^n$

Propiedad 5: Preservación de límites

Para cada $n = 1,2,3, \ldots$ sea $F(k/n)$ la Función de Distribución Acumulada (FDA) con su correspondiente FGM $M(\theta/n)$ definida sobre un intervalo abierto (θ_I, θ_S) que contiene a θ, para la cual:

$\lim_{n\to\infty} M\left(\frac{\theta}{n}\right) = M(\theta)$, para todo θ contenido en (θ_I, θ_S); entonces: (10.30)

$\lim_{n\to\infty} F(k/n) = F(x)$ para toda $x \in C = \{x|F(x)$ es continua en x$\}$ (10.31)

$F(X)$ es la FDA correspondiente a $M(\theta)$.

10.4 Funciones Características

Hemos visto que tanto la FGMF y la FGM no están definidas para todas las variables aleatorias; situación que no sucede con la Transformada de Fourier, conocida también como la Función Característica, la cual se define para las variables aleatorias discretas como

$$C_X(\theta) = \sum_{j=1}^{n} e^{i\theta x} p(x_i) \tag{10.32}$$

O para variables aleatorias continuas

$$C_X(\theta) = \int_{-\infty}^{\infty} e^{i\theta x} f(x) dx \tag{10.33}$$

En las cuáles $i = \sqrt{-1}$ es la unidad imaginaria, de donde se desprende que los valores de la función característica pueden ser número complejos y, para su estudio, se requiere del conocimiento de las variables complejas; por ello no se estudiarán aquí.

10.5 Momentos con respecto a un punto y la varianza

Los momentos que hasta ahora hemos estudiado, corresponden a los momentos respecto al origen -ver Figura 1.1-; Ahora generalizaremos estos conceptos estudiando los momentos con respecto a un punto c diferente del origen, también conocidos como momentos centrales para cuando $c = E[X]$.

Ya vimos como calcular el valor esperado de las variables aleatorias discretas y continuas respecto al origen; ahora ampliemos estos conceptos para el caso de la función particular $f(X) = (X - a)^r$, donde a y r son constantes, entonces a la expresión:

$$E[f(X)] = E[(X - a)^r] \tag{10.34}$$

se le conoce como el momento de orden *r* respecto a *a* y se representa con μ_r. Sebe tenerse presente las notaciones diferentes para los momentos factoriales -$\mu_{(r)}$-, los momentos ordinarios -μ'_r- y los momentos centrales -μ_r o $\mu_X^{(r)}$-; y, en particular, obsérvese que:

si $a = 0$, se tiene $E[(X-0)^r] = E[X^r] = \mu'_r$; para esta expresión, si $r = 1$ se tiene:

$\mu'_1 = E[X^1] = E[X]$, expresión que se conoce como *la media de la VA X,* concepto que se ampliará más adelante y, en este caso particular suele simbolizarse simplemente con μ, μ_X o μ'_1.

Ejemplo 10.27 Con referencia al Ejemplo 1, relacionado con la decisión que debe tomar el Director de PEMEX Exploración Producción si en determinada región debe o no perforar un pozo, evalúa esta situación mediante el valor esperado $E[X]$ o la media de la variable aleatoria X para las dos alternativas de decisión:

Si decide perforar: $\mu_X = E[X/I] = 200(0.25) - 50(0.75) = 12.50$
Si decide no perforar: $\mu_X = E[X/II] = 50(0.25) + 25(0.75) = 31.25$

Por lo tanto, a la luz de estos resultados de la *media*, la decisión debe ser No perforar.

Ejemplo 10.28 Con referencia al Ejemplo 10.2, relacionado con la proporción de impurezas de una muestra de petróleo crudo se modela por la función de densidad

$$f(x) = f(x) = \begin{cases} \frac{3}{2}x^2 + x, & 0 \leq x \leq 1 \\ 0 & otros\ casos \end{cases}$$

el valor esperado es la *media* de la proporción de impurezas:

$$\mu_X = E[X] = \int_0^1 x(\tfrac{3}{2}x^2 + x)dx = 0.71;$$

Que corresponde a la proporción de impurezas esperada.

Regresando a los momentos centrales, si:

- $r = 0$: $E[(X-a)^0] = E[1] = 1$;

- $a = E[X] = \mu'_1$:

$$\mu_r = \mu_X^{(r)} = E[(X - E[X])^r] = E[(X - \mu'_1)^r]$$

Que es la expresión generalizada de los momentos centrales μ_r, los cuáles pueden encontrarse a partir de los momentos ordinarios.

En efecto:

Si $r = 1$

$$\mu_1 = \mu_X^{(1)} = E[(X - E[X])^1] = E\big[X - E[X]\big] = E\big[X - E[X]\big] = E[X - \mu_1'],$$

y, por las propiedades de los valores esperados:

$$\mu_1 = \mu_X^{(1)} = E[X] - E[\mu_1'] = \mu_1' - \mu_1' = 0;$$

Cuando $r = 2$**:**

$$\boldsymbol{\sigma_X^2 = \mu_2 = \mu_X^{(2)} = Var[X] = V[X] = E[(X - \mu_1')^2]} \tag{10.35}$$

Esta expresión se conoce como *la varianza de la VA X,* que se simboliza con $\boldsymbol{\sigma_X^2, V[X], Var[X], \mu_2}$ **o** $\boldsymbol{\mu_X^{(2)}}$**, y es la desviación esperada al cuadrado de** $\boldsymbol{X}$ **con respecto a la media, o el promedio de las desviaciones con respecto a la media; constituye uno de los parámetros de dispersión de dicha variable porque es la desviación más pequeña que se obtiene, cuando se calcula a partir de la media** $\boldsymbol{\mu_1'}$**.**

En efecto, si se elige un valor arbitrario a y calculamos una pseudo varianza

$$\sigma_{X/a}^2 = \mu_{X/a}^{(2)} = E[(X-a)^2] = E[(X - \mu_X + \mu_X - a)^2] = E\{[(X - \mu_X) + (\mu_X - a)]^2\}$$

$$= E[(X - \mu_X)^2 + 2(X - \mu_X)(\mu_X - a) + (\mu_X - a)^2]$$

$$= E[(X - \mu_X)^2] + 2(\mu_X - a)E[(X - \mu_X)] + (\mu_X - a)^2$$

$$= E[(X - \mu_X)^2] + 0 + (\mu_X - a)^2$$

$$\sigma_{X/a}^2 = \sigma_X^2 + (\mu_X - a)^2$$

Por lo que $\sigma_{X/a}^2 \geq \sigma_X^2 + (\mu_X - a)^2$.

Y cuando $a = \mu_X$ se tiene $\sigma_{X/a}^2 = \sigma_X^2$

Lo que demuestra que la varianza respecto a la media siempre es más pequeña que la pseudo varianza respecto a cualquier otro punto a.

10.6 Álgebra de las Varianzas

Al igual que para los valores esperados, es de mucha utilidad establecer algunas reglas para la varianza que se aplican a las variables aleatorias discretas y continuas, a partir de la aplicación del valor esperado.

Regla 1. La varianza en términos de los valores esperados

La varianza de una variable aleatoria es igual al segundo momento ordinario respecto al origen al cuadrado menos el cuadrado del primer momento ordinario respecto al origen; es decir, es igual al valor esperado del cuadrado de la variable aleatoria menos el cuadrado del valor esperado de la variable aleatoria.

$$\sigma_X^2 = \mu_2 = \mu_X^{(2)} = \mu_2' - (\mu_1')^2 = E[X^2] - E[X]^2 \tag{10.36}$$

En efecto:

$$\sigma_X^2 = E[(X - \mu_1')^2] = E\left[X^2 - 2X\mu_1' + {\mu_1'}^2\right]$$

$$\sigma_X^2 = E[X^2] - 2\mu_1' E[X] + E[\mu_1']^2 = E[X^2] - 2(\mu_1')^2 + (\mu_1')^2$$

$$\sigma_X^2 = \mu_2 = \mu_X^{(2)} = \mu_2' - (\mu_1')^2 = E[X^2] - E[X]^2$$

Regla 2. La Varianza de una constante

Como una constante no varía, su varianza es cero; es decir si $a \in \mathbb{R}$, entonces:

$$\sigma_a^2 = V(a) = 0 \tag{10.37}$$

En efecto:

$$\sigma_a^2 = E[a^2] - E[a]^2 = a^2 - a^2 = 0$$

Regla 3. La Varianza de una constante por una variable aleatoria

Si $a \in \mathbb{R}$ y X es una variable aleatoria con varianza σ_X^2, entonces la varianza de la variable aleatoria aX es

$$\sigma_{aX}^2 = V(aX) = a^2\sigma_X^2 \tag{10.38}$$

En efecto:

$$\sigma_{aX}^2 = E[(aX)^2] - E[aX]^2 = a^2E[X^2] - a^2E[X]^2 = a^2(E[X^2] - E[X]^2)$$

$$\sigma_{aX}^2 = a^2\sigma_X^2$$

Lo que significa que al multiplicar cada valor de la variable aleatoria por una constante se multiplica la varianza por el cuadrado de la constante, y a actúa como un factor de escala y por lo tanto un cambio en la dispresión.

Regla 4. La Varianza de una combinación lineal

Si $a \in \mathbb{R}$ y X es una variable aleatoria con valor esperado $E[X]$ y varianza σ_X^2, entonces la varianza de la variable aleatoria $X + a$ es:

$$\sigma_{X+a}^2 = V(X + a) = V(X) = \sigma_X^2 \tag{10.39}$$

En efecto:

$$\sigma^2_{X+a} = V(X+a) = E[(X+a) - E[X+a]]^2 = E[(X+a) - (E[X]+a)]^2$$

$$\sigma^2_{X+a} = V(X+a) = E[X + a - E[X] - a]^2 = E\big[X - E[X]\big]^2$$

$$\sigma^2_{X+a} = \sigma^2_X$$

Lo que significa que al añadir una constante a cada valor de la variable aleatoria su varianza permanece sin cambio y a *actúa como una constante de desplazamiento* de la distribución de la variable aleatoria a lo largo del eje *x.*

Ahora bien, Para $r = 3$ tenemos:

$$\mu_X^{(3)} = \mu_3 = E[(X - \mu_1')^3] = E\big[X^3 - 3X^2\mu_1' + 3X{\mu_1'}^2 - {\mu_1'}^3\big]$$

Aplicando las propiedades del valor esperado:

$$\mu_X^{(3)} = \mu_3 = E[X]^3 - 3\mu_1' E[X^2] + 3{\mu_1'}^2 E[X] - {\mu_1'}^3$$

$$\mu_X^{(3)} = \mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2{\mu_1'}^3$$

Siguiendo el mismo procedimiento para $r = 4$ se obtiene:

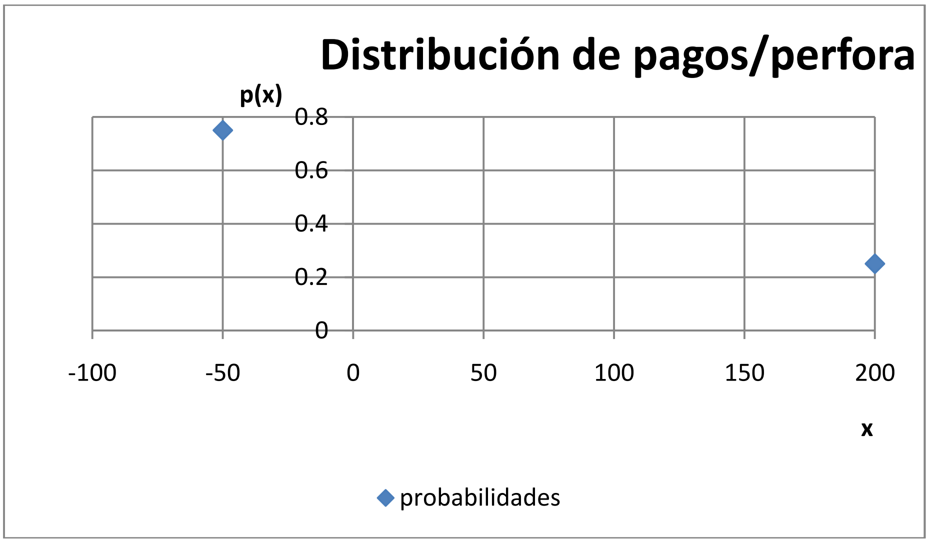
$$\mu_X^{(4)} = \mu_4 = \mu_4' - 4\mu_1'{\mu_3'}^2 + 6\mu_1'{\mu_2'}^2 - 3{\mu_1'}^4$$

Ejemplo 10.29 Para Ejemplo 10.1 calculemos los cuatro momentos centrales, recordando que el segundo momento central es la varianza.

Con esta información se obtiene la siguiente tabla de pagos, comúnmente utilizada en la teoría de decisiones:

Tabla 10.1. matriz de pagos

		Estado de la	**naturaleza (X)**
Decisión por tomar		I: SÍ existe petróleo	II: NO existe petróleo
	Sí perforar	200	-50
	No perforar	50	25



Para evaluar esta situación, se aplica el valor esperado para las dos alternativas de decisión:

Si decide perforar: $E[X/\mathrm{I}] = 200(0.25) - 50(0.75) = 12.50$

Si decide no perforar: $E[X/\mathrm{II}] = 50(0.25) + 25(0.75) = 31.25$

E[X/I]	E[X/II]	E[X^2/I]	E[X^2/II]	E[X^3/I]	E[X^3/II]	E[X^4/I]	E[X^4/II]
12.5		1187		1906250		404687500	
	31.25		1093.75		42968.7		1855468.75

$\mu_{2/I}$	$\mu_{2/II}$	$\mu_{3/I}$	$\mu_{3/II}$	$\mu_{4/I}$	$\mu_{4/II}$
11718.75		1464843.75		-19818627930	
	117.1875		1464.84375		32043.457

Ejemplo 10.30 Para el Ejemplo 10.2 La proporción de impurezas de una muestra de petróleo crudo se modela por la función de densidad:

$$f(x) = f(x) = \begin{cases} \frac{3}{2}x^2 + x, & 0 \leq x \leq 1 \\ 0 & otros\ casos \end{cases}$$

$$E[X] = \int_0^1 x(\tfrac{3}{2}x^2 + x)dx = \int_0^1 (\tfrac{3}{2}x^3 + x^2)dx = \left[\frac{3}{2} \times \frac{x^4}{4} + \frac{x^3}{3}\right]_0^1 = \frac{3}{8} + \frac{1}{3} = \frac{17}{24} = 0.71;$$

$$E[X^2] = \int_0^1 x^2(\tfrac{3}{2}x^2 + x)dx = \int_0^1 (\tfrac{3}{2}x^4 + x^3)dx = [\frac{3x^5}{10} + \frac{x^4}{4}]_0^1 = \frac{3}{10} + \frac{1}{4} = \frac{11}{20} = 0.55$$

$$E[X^3] = \int_0^1 x^3(\tfrac{3}{2}x^2 + x)dx = \int_0^1(\tfrac{3}{2}x^5 + x^4)dx = [\tfrac{3x^6}{12} + \tfrac{x^5}{5}]_0^1 = \tfrac{3}{12} + \tfrac{1}{5} = \tfrac{9}{20} = 0.45$$

$$E[X^4] = \int_0^1 x^4(\tfrac{3}{2}x^2 + x)dx = \int_0^1(\tfrac{3}{2}x^6 + x^5)dx = [\tfrac{3x^7}{14} + \tfrac{x^6}{6}]_0^1 = \tfrac{3}{14} + \tfrac{1}{6} = \tfrac{8}{21} = 0.38$$

$$\mu_1 = \mu_X^{(1)} = E[X] - E[\mu_1'] = \mu_1' - \mu_1' = 0;$$

$$\sigma_X^2 = \mu_2 = \mu_X^{(2)} = \mu_2' - (\mu_1')^2 = E[X^2] - E[X]^2 = 0.55 - (0.71)^2 = .046$$

$$\begin{aligned}\mu_3 = \mu_X^{(3)} &= \mu_3' - 3\mu_1'\mu_2' + 2{\mu_1'}^3 = E[X^3] - 3E[X]E[X^2] + 2E[X]^3 \\ &= 0.45 - 3(0.71)(0.55) + 2(0.45)^3 = -0.32\end{aligned}$$

$$\begin{aligned}\mu_4 = \mu_X^{(4)} &= \mu_4' - 4\mu_1'{\mu_3'}^2 + 6\mu_1'{\mu_2'}^2 - 3{\mu_1'}^4 \\ &= E[X^4] - 4E[X]E[X^3]^2 + 6E[X]E[X^2]^2 - 3E[X]^4 \\ &= 0.38 - 4(0.71)(0.45)^2 + 6(0.71)(0.55)^2 - 3(0.71)^4 = 0.33\end{aligned}$$

La potencialidad de los momentos se verá con claridad cuando estudiemos las distribuciones de probabilidad de las variables aleatorias continuas y discretas.

10.7 Momentos de Distribuciones condicionales y conjuntas

Los momentos de las distribuciones condicionales se basan en los conceptos de valor esperado y de la distribución condicional.

Para las VA discretas

$$E[X|y] = \textstyle\sum_{\forall x} xp(x|y) \quad \text{o} \quad E[Y|x] = \sum_{\forall y} yp(y|x) \tag{10.40}$$

Para las VA continuas

$$E[X|y] = \int_{-\infty}^{\infty} xf(x|y)dx \quad \text{o} \quad E[Y|x] = \int_{-\infty}^{\infty} yf(y|x)dx \tag{10.41}$$

Y, por el concepto de variables aleatorias independientes, se tiene que si X y Y son independientes se deberá cumplir que

$$E[X|y] = E[X] \qquad \text{o} \qquad E[Y|x] = E[Y] \tag{10.42}$$

Este concepto de momentos puede extenderse a las funciones conjuntas de dos o más variables; SI X Y Y son VA conjuntas con función de probabilidad conjunta $P(x,y)$ para el caso discreto o función de densidad $f(x,y)$ para el caso continuo y si $g(X,Y)$ es una función de X Y Y, se tiene

Para el caso discreto

$$E[g(X,Y)] = \textstyle\sum_{\forall x}\sum_{\forall y} g(x,y)P(x,y) \tag{10.43}$$

Y para el caso continuo

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f(x,y) dydx \tag{10.44}$$

Así, para una distribución conjunta de dos variables aleatorias X y Y, para m y n enteros no negativos los valores esperados de productos de potencias son $E[X^m y^n]$

En particular, si X Y Y son independientes

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dydx = \int_{-\infty}^{\infty} xf(x)dx \int_{-\infty}^{\infty} yf(y)dy = E[X]E[Y] \tag{10.45}$$

Y por extensión para el caso de X_i, $i = 1,2,3, \dots, n$ VA independientes se tiene

$$E[X_1X_2X_3 \dots X_n] = E[X_1]E[X_2]E[X_3] \dots E[X_n] \tag{10.46}$$

10.8 La covarianza

El momento de la distribución conjunta de las variables X Y Y que refleje la fuerza de la dirección de la relación es la covarianza, que se define como

$$\sigma_{XY} = COV[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] \tag{10.47}$$

Cuando X Y Y son discretas se tiene

$$\sigma_{XY} = \sum_{\forall x}\sum_{\forall y}(x-\mu_X)(y-\mu_Y)\, p(X=x, Y=y) \tag{10.48}$$

Y si son continuas

$$\sigma_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)\, f(x,y)dydx \tag{10.49}$$

Como $E[(X-\mu_X)(Y-\mu_Y)] = E[XY - \mu_Y X - \mu_X Y + \mu_X\mu_Y]$

Por las propiedades de los valores esperados, se tiene

$$\sigma_{XY} = COV[X,Y] = E[XY] - E[X]E[Y] \tag{10.50}$$

Cuando X Y Y son discretas se tiene

$$\sigma_{XY} = \sum_{\forall x}\sum_{\forall y} xy\, p(X=x, Y=y) - \sum_{\forall x} x\, p(X=x) \sum_{\forall y} y\, p(Y=y) \tag{10.51}$$

Y para cuando son continuas

$$\sigma_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)dydx - \int_{-\infty}^{\infty} xf(x)dx \int_{-\infty}^{\infty} yf(y)dy \tag{10.52}$$

Y si X Y Y son independientes

$$\sigma_{XY} = COV[X,Y] = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0 \qquad (10.53)$$

Lo que significa que la dirección de la interrelación de las variables X Y Y no tiene ningún patrón positivo o negativo. Si X Y Y están positivamente relacionadas entonces valores mayores de X tienden a corresponder con valores mayores de Y y la $COV[X,Y]$ es positiva**;** en cambio**, si** valores menores de X se corresponden con valores menores de Y se tiene que la $COV[X,Y]$ es negativa**. En suma, el signo de la covarianza nos proporciona información a cerca de la dirección de la interrelación existente entre** X Y Y.

La covarianza es particularmente útil en la determinación de la suma, de la diferencia o el coeficiente de correlación de dos variables aleatorias. Así, si $Z = X + Y$,

Conforme a (10.35) se tiene

$$V[Z] = V[X+Y] = E[(X+Y) - E[X+Y]]^2$$

Como $E[X+Y] = E[X] + E[Y] = \mu_X + \mu_Y$, se tiene

$$V[Z] = V[X+Y] = E[(X+Y) - (\mu_X + \mu_Y)]^2 = V[X+Y] = E[(X-\mu_X) + (Y-\mu_Y)]^2$$

$$V[Z] = V[X+Y] = E[(X-\mu_X)^2 + (Y-\mu_Y)^2 + 2(X-\mu_X)(Y-\mu_Y)]$$

Por las propiedades de valor esperado

$$V[Z] = V[X+Y] = E[(X-\mu_X)^2] + E[(Y-\mu_Y)^2] + 2E[(X-\mu_X)(Y-\mu_Y)]$$

$$V[Z] = V[X+Y] = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY} \qquad (10.54)$$

Si ahora $W = X - Y$, por el mismo procedimiento se tiene

$$V[W] = V[X-Y] = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY} \qquad (10.55)$$

Y, si las VA son independientes se tiene

$$V[Z] = V[X+Y] = V[W] = V[X-Y] = \sigma_X^2 + \sigma_Y^2 \qquad (10.56)$$

Es conveniente observar la influencia de la covarianza sobre la varianza de la suma o diferencia de las variables aleatorias, salvo cuando son independientes; si la relación es positiva la varianza de la suma es mayor y la varianza de la diferencia es menor que si fueran independientes. En contraparte si la relación entre X y Y es negativa, se tiene que la varianza de la suma es menor y la varianza de la diferencia es mayor que si las variables que si ellas fueran independientes.

10.9 El Coeficiente de Correlación

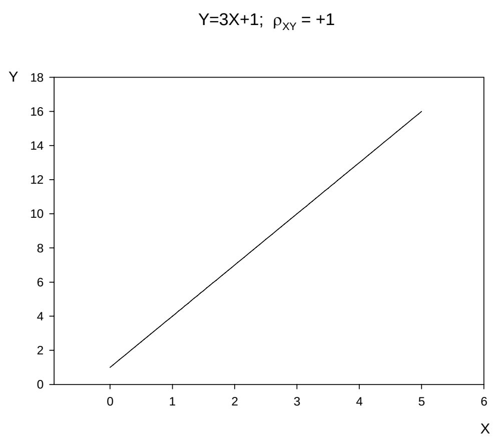
Si se observa, la covarianza tiene como unidades las correspondientes a X Y Y; por ejemplo si X está medida en metros y Y en kg sus unidades son m-kg y es difícil de interpretar la fuerza de la relación entre X Y Y. Esta dificultad se resuelve dividiendo la covarianza por el producto de las desviaciones estándar de estas variables σ_X y σ_Y dando por resultado un parámetro adimensional, que es uno de los parámetros generales de las Funciones de probabilidad conjuntas, y se conoce como **el coeficiente de correlación y se define como**

$$\boldsymbol{\rho_{XY}} = \boldsymbol{\rho} = \frac{\boldsymbol{COV[X,Y]}}{\boldsymbol{\sigma_X \sigma_Y}} = \frac{\sigma_{XY}}{\boldsymbol{\sigma_X \sigma_Y}} \tag{10.57}$$

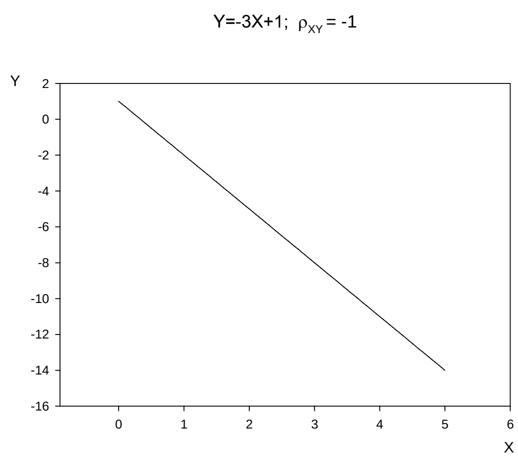
Puede demostrarse que $-1 \leq \rho_{XY} \leq +1$

Así pues, el coeficiente de correlación corrige el escalamiento de X Y Y y representa apropiadamente la fuerza de y la dirección de la relación que existe entre dichas variables. Si $\rho_{XY} = +1$ los valores de las variables están perfectamente correlacionadas positivamente como se muestra en la figura 10.2 (a); en cambio para $\rho_{XY} = -1$ los valores de las variables estarán perfectamente correlacionadas negativamente como se indica en la figura 10.2 (b); y si $\rho_{XY} = 0$ se dice que las variables NO están correlacionadas como se ilustra en (c) de la figura. Para los demás valores si ρ_{XY} está cercano a $+1$ o a -1 se tiene una interrelación fuerte, en cambio si está cercano a 0 implica una relación muy débil.

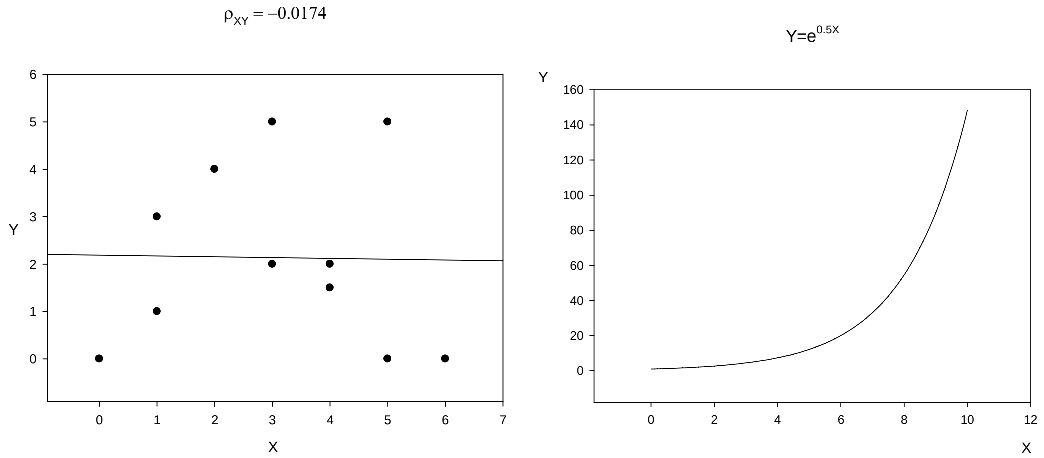
En necesario tener en cuenta que el coeficiente de correlación $\boldsymbol{\rho_{XY}}$ **mide la fuerza de la relación lineal entre las variables bajo estudio** $\boldsymbol{X}$ **Y** $\boldsymbol{Y}$; por ejemplo, si $Y = 3X + 1$ las variables están perfectamente correlacionadas positivamente, en cambio para $Y = e^{0.5X}$ las variables no están perfectamente relacionadas porque la relación no es lineal, como se ilustra en la figura 10.2.



(a)



(b)



(c) (d)

Figura 10.2 Relaciones lineal y no lineal

Ejemplo 10.31 Si X es el número de libros de Teoría de la Probabilidad que se tienen en una biblioteca y Y el de libros de Estadística Aplicada, y la distribución conjunta de la selección de un estudiante es la que aparece en el cuadro interior con la cual se tiene lo siguiente:

a) Las distribuciones marginales de X y Y que son los valores de las variables conjuntamente con las probabilidades que aparecen en el margen derecho -para Y- y en el inferior -para X-.

		X			
		1	2	3	p(y)
	1	0.000	0.167	0.083	0.250
Y	2	0.200	0.111	0.000	0.311
	3	0.133	0.250	0.056	0.439
p(x)		0.333	0.528	0.139	1.000

b) La probabilidad $p(Y = 3|X = 2) = \frac{p(Y=3,X=2)}{p(X=2)} = \frac{0.250}{0.528} = 0.474$

c) La covarianza de X y $\mathrm{Y} = COV[X, Y] = \mathrm{E}[\mathrm{X}, \mathrm{Y}] - \mu_X \mu_Y$

$$E[X, Y] = (1)(1)(0) + (1)(2)(0.167) + (1)(3)(\ 0.083) + \cdots + (3)(3)(\ 0.056)$$
$$= 3.828$$

$$\mu_X = 1(0.333) + 2(0.528) + 3(0.139) = 1.806$$

Con el mismo procedimiento $\mu_Y = 2.189$.

Por lo tanto $COV[X, Y] = 3.828 - 1.806(2.189) = -0.124$

d) Para calcular el coeficiente de correlación necesitamos calculas las desviaciones estándar de las variables, lo que a su vez requiere del cálculo de las varianzas.

$$E[X^2] = 1^2(0.333) + 2^2(0.528) + 3^2(0.139) = 3.694$$

$$Var[X] = E[X^2] - \mu_X{}^2 = 3.694 - 1.806^2 = 0.434$$

Y la desviación estándar es $\sigma_X = \sqrt{0.434} = 0.659$

Siguiendo el mismo procedimiento se tiene $E[Y^2] = 5.444; \sigma_Y^2 = 0.653$ yσ_Y= 0.808

Con lo cual tenemos lo necesario para calcular el coeficiente de correlación de X y Y

$$\rho_{XY} = \frac{COV[X,Y]}{\sigma_X \sigma_Y} = \frac{-0.124}{(0.659)(0.808)} = -0.233$$

Lo que significa que existe muy poca relación lineal negativa y significa que cuando se eligen libros de la Teoría de la probabilidad se tiene una menor elección de libros de Estadística Aplicada.

Ejemplo 10.32 Sean X y Y dos variables aleatorias con función de densidad conjunta

$$f(x,y) = \begin{cases} k(x^2+y^2) & para\ 0 < x < 2; 1 < y < 4 \\ 0 & para\ otros\ valores \end{cases}$$

a) Calculemos el valor de k.

$\int_0^2 \int_1^4 k(x^2+y^2)dydx$ debe ser igual a 1.

$$\int_0^2 \int_1^4 k(x^2+y^2)\,dydx = k\int_0^2 (x^2y + \frac{y^3}{3})\Big|_{y=1}^{y=4} dx = k\int_0^2 (3x^2+21)dx$$

$= 50k = 1 \ \therefore k = \frac{1}{50}$

Con lo cual

$$f(x,y) = \begin{cases} \frac{1}{50}(x^2+y^2) & para\ 0 < x < 2; 1 < y < 4 \\ 0 & para\ otros\ valores \end{cases}$$

b) Calculemos el coeficiente de correlación.

Como $\rho_{XY} = \frac{COV[X,Y]}{\sigma_X \sigma_Y} = \frac{\mathrm{E[X,Y]} - \mu_X \mu_Y}{\sigma_X \sigma_Y}$

I. Con las densidades marginales podemos calcular las medias y las desviaciones estándar de X y Y.

La FDM de X es

$$f(x) = \int_1^4 \frac{1}{50}(x^2 + y^2)dy = \frac{1}{50}(x^2 y + \frac{y^3}{3})\Bigg|_{y=1}^{y=4} = \frac{1}{50}(3x^2 + 21)$$

Calculemos la media de X

$$\mu_X = \frac{1}{50}\int_0^2 x(3x^2 + 21)\,dx = \frac{1}{50}(12 + 42) = 1.08$$

Ahora $E[X^2]$

$$E[X^2] = \frac{1}{50}\int_0^2 x^2(3x^2 + 21)\,dx = \frac{1}{50}\int_0^2 (3x^4 + 21x^2)\,dx = 2.08$$

La varianza vale $Var[X] = E[X^2] - {\mu_X}^2$

$$Var[X] = 2.08 - 1.08^2 = 0.9136$$

Y $\sigma_X = \sqrt{0.9136} = 0.9558$

II. Siguiendo los mismos procedimientos para la variable aleatoria Y tenemos

La FDM de Y es $\int_0^2 \frac{1}{50}(x^2 + y^2)dx = \frac{1}{50}(y^2 x + \frac{x^3}{3})\Big|_{x=0}^{x=1} = \frac{1}{50}(2y^2 + \frac{8}{3})$

Calculemos la media de Y

$$\mu_Y = \frac{1}{50}\int_1^4 y(2y^2 + \frac{8}{3})\,dx = \frac{1}{50}(147.5) = 2.95$$

Ahora $E[Y^2]$

$$E[Y^2] = \frac{1}{50}\int_1^4 y^2(2y^2 + \frac{8}{3})\,dx = \frac{1}{50}\int_0^2 (2y^4 + \frac{8}{3}y^2)\,dy = 9.304$$

La varianza vale $Var[Y] = E[Y^2] - {\mu_Y}^2$

$$Var[Y] = 9.304 - 2.95^2 = 0.6015$$

Y la desviación estándar

$\sigma_Y = \sqrt{0.6015} = 0.7755$

III. La varianza conjunta vale

IV. $\sigma_{XY} = E[XY] = \int_0^2 \int_1^4 xy[\frac{1}{50}(x^2 + y^2)]dydx = \frac{1}{50}\int_0^2 \int_1^4 (x^3y +$ *xy3dydx=150027.5x3+63.75xdx=3.15*

Con los valores anteriores tenemos

V. $\rho_{XY} = \frac{\mathrm{E[X,Y]} - \mu_X\mu_Y}{\sigma_X\sigma_Y} = \frac{3.15-1.08(2.95)}{0.9558(0.7755)} = -0.0485$

Lo que indica que prácticamente no existe correlación lineal entre las variables aleatorias X y Y.

10.10 La curva de regresión y el Coeficiente de Determinación

Como se vio en la sección anterior, el coeficiente de correlación da una medida de la fuerza de asociación lineal entre las variables aleatorias X y Y; por lo tanto, si dicho coeficiente de correlación es satisfactorio, entonces es deseable buscar un procedimiento para poder pronosticar el valor de Y -suponiendo que ésta es la variable dependiente- en términos de los valores x de X -que para nuestro caso es la variable independiente-. Puesto que x es un valor conocido, para la predicción de Y se utiliza la distribución condicional de Y dado el valor x de X; es decir, $p(Y|x)$ o $f(y|x)$ según se trate de una VA discreta o continua, cuyo valor esperado $E[Y|x] = \mu_{Y|x}$ es el mejor estimador o el más razonable predictor de Y.

Como el valor esperado $E[Y|x] = \mu_{Y|x}$ varía para los diferentes valores de X, la función formada por las parejas de valores $\{x, E[Y|x]\}$ se llama la curva de regresión de Y $sobre$ x; la cual se representa en la figura 10.3.

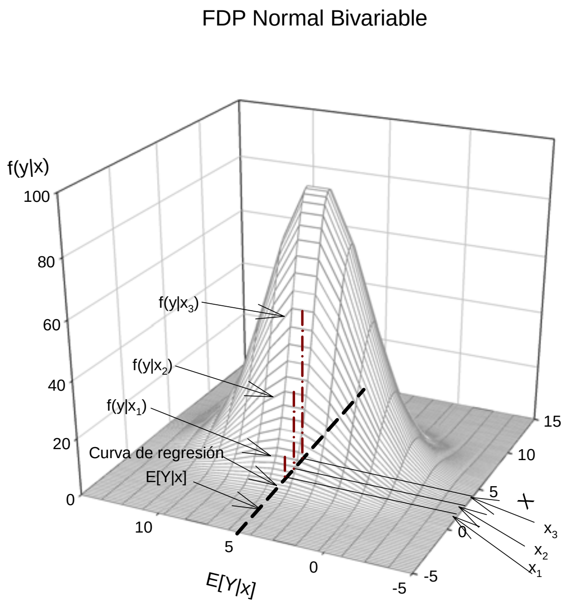


Figura 10.3 Curva de Regresión E[Y|X=x]

En este caso muy particular $E[Y|x]=5$; lo que significa que Y y X son independientes o sea que X no predice a Y, no obstante, las siguiente explicaciones son válidas. En la figura se observa que la superficie corresponde a una distribución conjunta normal bivariable generada por un número infinito de FDP condicionales $f(y|x)$ -una para cada valor de X; para este caso particular la curva de regresión es una recta de regresión sobre el plano X-E[Y|x]; esta curva se encuentra graficando las parejas de valores $\{x, E[Y|x]\}$, donde $E[Y|x]$ son los valores esperados o las medias de la FDP condicionales $f(y|x_i)$; se han dibujado solamente 3 valores de X para los cuales se tienen las 3 FDP condicionales $f(y|x_i) para\ \ i=1,2,3$; y cada una tiene su valor esperado $E[Y|x_i]$ sobre el plano $X-E[Y|x]$; la infinidad de estos puntos $\{x, E[Y|x]\}$ define la curva de regresión como se ilustra en la figura 10.4.

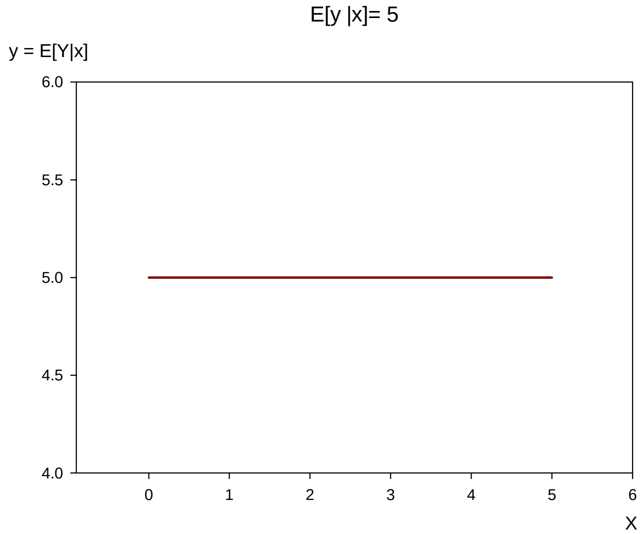


Figura 10.4 la Recta de Regresión $E[Y|x]=5$

Cabe recordar que Y es la variable dependiente de X -la variable independiente- y la curva de regresión de Y sobre x es diferente a la curva de regresión de X sobre y.

Ejemplo 10.33 Con relación al ejemplo anterior determinemos la curva de regresión $E[Y|x]$. Sabemos que

$$f(x,y)=\begin{cases}\frac{1}{50}(x^2+y^2) & para\ 0<x<2; 1<y<4\\ 0 & para\ otros\ valores\end{cases}$$

Y encontramos que

$$f(x)=\frac{1}{50}(3x^2+21)\ \ para\ 0<x<2$$

Con estas FDP encontremos la FDP condicional $f(y|x)$.

$$f(y|x)=\frac{f(x,y)}{f(x)}=\frac{\frac{1}{50}(x^2+y^2)}{\frac{1}{50}(3x^2+21)}=\frac{(x^2+y^2)}{(3x^2+21)}$$

Como ya vimos, la curva de regresión de Y sobre x es el valor esperado de la FDP condicional $f(y|x)$.

$$E[Y|x] = \int_1^4 y\, f(y|x)dy = \int_1^4 \frac{(yx^2 + y^3)}{(3x^2 + 21)} dy$$

$$= \frac{1}{(3x^2+21)} \int_1^4 (yx^2 + y^3)dy = \frac{1}{(3x^2+21)} \left[\frac{y^2x^2}{2} + \frac{y^4}{4}\right] \Bigg|\begin{matrix} y=4 \\ y=1 \end{matrix} = \frac{7.5x^2 + 63.75}{3x^2 + 21}$$

La curva de regresión se muestra en la figura 10.5

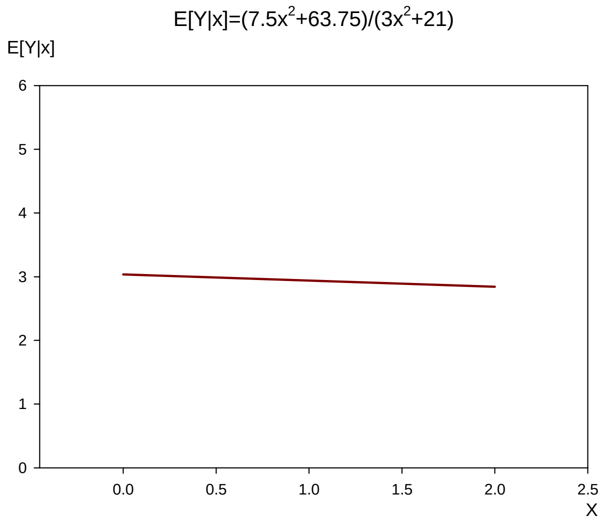


Figura 10.5 Curva de regresión de Y sobre x: $E[Y|x]$

El cálculo de la curva de regresión de $E[X|y]$ se hace siguiendo el mismo procedimiento.

$$f(x|y) = \frac{f(x,y)}{f(y)} = \frac{\frac{1}{50}(x^2+y^2)}{\frac{1}{50}(2y^2+\frac{8}{3})} = \frac{(x^2+y^2)}{\left(2y^2+\frac{8}{3}\right)}$$

$$E[X|y] = \int_0^2 x\, f(x|y)dx = \int_0^2 \frac{x(x^2+y^2)}{\left(2y^2+\frac{8}{3}\right)} dx$$

$$= \frac{1}{\left(2y^2+\frac{8}{3}\right)} \int_0^2 (xy^2 + x^3)dx = \frac{1}{\left(2y^2+\frac{8}{3}\right)} \left(\frac{x^2y^2}{2} + \frac{x^4}{4}\right)\Bigg|\begin{matrix} x=2 \\ x=0 \end{matrix}$$

$$E[X|y] = \frac{2y^2+4}{2y^2+\frac{8}{3}}$$

La representación gráfica de esta curva de regresión aparece en la figura 10.6 y, similar a la anterior, no es útil para predecir los valores de X dado los diferentes valores que Y puede tomar.

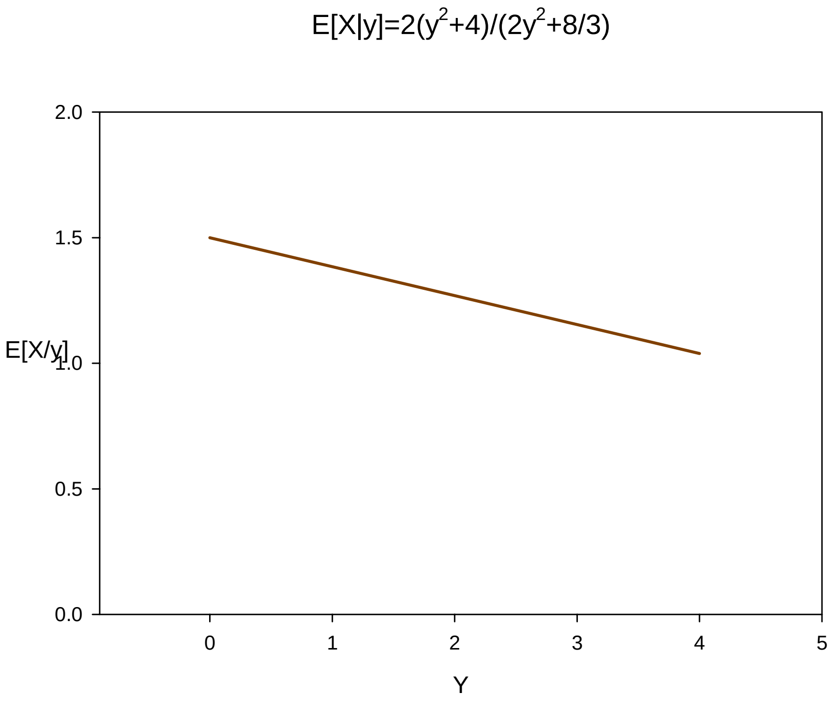


Figura 10.6 Curva de regresión X sobre $E[X|y]$

10.11 El Coeficiente de Determinación

El coeficiente de correlación nos da un valor de la fuerza de asociación lineal entre las variables Y y X, en tanto que la curva de regresión es el modelo matemático o la función de la relación entre las variables, que puede o no ser lineal, y otro parámetro que relaciona tanto a la potencia de asociación como a la regresión es el coeficiente de determinación que nos da una medida de la varianza explicada por la curva de regresión y se define como el cuadrado del coeficiente de correlación.

$$COEFICENTE\ DE\ DETERMINACIÓN = \rho_{XY}{}^2 \qquad (10.58)$$

Ejemplo 10.34 Como era de esperarse, las curvas de regresión del ejemplo anterior no son útiles para predecir los valores de Y con los diferentes valores que X puede tomar, ni los de X los valores de Y; ya que, del ejemplo anterior, el coeficiente de correlación ρ_{XY} resultó igual a -0.0485 y el coeficiente de determinación es

$$\rho_{XY}{}^2 = (-0.0485)^2 = 0.0024$$

Lo que significa que el 0.24% de la varianza de Y o de X es explicada por las curvas de regresión.

En la parte de estadística aplicaremos estos conceptos a los datos de las muestras y dedicaremos un capítulo a ellos.

10.12 Bibliografía y Referencias

Lindgrend B. (1968), *statistical theory,* Second Edition, THE MACMILLAN COMPANY, USA.

Olkin I. Gleser L. Derman C. (1980), *PROBABILITY MODELS AND APPLICATIONS*, Macmillan Publishing, USA.

Winkler R. Hays W. (1971), ***Statistics**, probability, inference, and decision*, second edition, HOLT, RINEHART AND WINSTON, USA.

Devore J. (2005), *PROBABILIDAD Y ESTADÍSTICA para ingeniería y ciencias*, International Thomson Editores, México.

Wackerly D. Mendenhall W. Scheaffer R. (2002), *ESTADÍSTICA MATEMÁTICA con aplicaciones*, sexta edición, International Thomson Editores, México.

Johnson N. Leone F. (1977), ***Statistics and Experimental design**, in Engineering and Physical Sciences*, Vol. I, Second Edition, John Wiley &. USA.

Hines W. Montgomery D. (1980), *PROBABILITY AND STATISTICS IN ENGINEERING AND MANAGEMENT SCIENCE*, 2nd. ED., John Wiley &Sons, Canada.

Nombre de archivo:	cap10 VALORES ESPERADOS Y MOMENTOS (definitivo).docx
Directorio:	C:\Documents and Settings\bfc\Mis documentos\g-1) Capítulos de mi libro de Probabilidad y estadística 2007-2009
Plantilla:	C:\Documents and Settings\bfc\Datos de programa\Microsoft\Plantillas\Normal.dotm
Título:	
Asunto:	
Autor:	fACULTAD DE INGENIERÍA
Palabras clave:	
Comentarios:	
Fecha de creación:	20/08/2009 8:44:00
Cambio número:	22
Guardado el:	08/02/2010 22:54:00
Guardado por:	fACULTAD DE INGENIERÍA
Tiempo de edición:	1,103 minutos
Impreso el:	30/08/2010 13:06:00
Última impresión completa	
Número de páginas:	34
Número de palabras:	8,721 (aprox.)
Número de caracteres:	47,966 (aprox.)