I.Q.A. LUIS OMAR JAMED BOZA
Dr. Rafael Lucio No. 199-D
Xalapa Eqz., Ver., México C.P. 91000
Tel. (01)(228)8155385
Cel. (044228)9794615
lojb33@yahoo.com.mx
lojb@prodigy.net.mx

Tema A04
Temas suplementarios para reforzar las materias Estadísticas

CONTENIDO GENERAL

4. Distribuciones Conjuntas
4.1 Consideraciones Generales de Distribuciones Conjuntas ---------- 2
4.2 Distribuciones Conjuntas Marginales ---------- 2
4.3 Distribuciones Conjuntas Condicionales ---------- 6
5. Distribuciones de muestreos que se derivan de la Distribución Normal
5.1 Chi–Cuadrada ---------- 6
5.2 t–Student ---------- 7
5.3 F de Fisher ---------- 7
6. Anexos a los Temas (1, 2, 3, 4 y 5)
6.1 Concepto de Independencia en las Variables Aleatorias ---------- 9
6.2 Concepto de Covarianza y Correlación ---------- 9

🡐

4. Distribuciones Conjuntas

Consideraciones Generales de Distribuciones Conjuntas

En muchos fenómenos en la naturaleza, es necesario tratar con más de una variable para modelizar el comportamiento del fenómeno, en éstos casos, debemos trabajar con dos o más variables aleatorias simultáneamente, teniendo un comportamiento conjunto dependiendo de los valores de las variables involucradas.

Si $[X_1, X_2, ..., X_K]$ es un vector aleatorio de funciones K-dimensionales, asociado a un espacio de resultados **"ε"**, entonces el Rango **"R"** de valores $[x_1, x_2, ..., x_K]$ que pueden tener resultado es un producto cartesiano de los conjuntos del espacio del rango de los componentes **(k = 1, 2, ...,K)**, es decir, lo podemos expresar de la siguiente forma:

$$R = \{[x_1, x_2, ..., x_K] : x_1 \in R_{X1}, x_2 \in R_{X2}, ... x_K \in R_{XK}\}$$

El espacio de resultados **"ε"**, debe quedar definido para los casos en que el vector anterior sea Discreto en sus "K" variables, o sea Continuo en sus "K" variables o el caso en donde haya de las 2 formas, éste último no lo tocaré en ésta síntesis.

Caso en que las "K" variables son Discretas

La Probabilidad del evento **"ε"** conjunto, debe cumplir lo siguiente:

1. $P(\varepsilon) = P(x_1, x_2, ..., x_K) \geq 0$ para toda **(k = 1, 2, ...,K)**
2. $\sum_{\text{todo 1}} \sum_{\text{todo 2}} ..... \sum_{\text{todo k}} P(x_1, x_2, ..., x_K) \equiv 1$
3. Los valores $([x_1, x_2, ..., x_K], P(x_1, x_2, ..., x_K))$ para toda **(k = 1, 2, ...,K)** forman la **Distribución de Probabilidad Conjunta**, del vector original $[X_1, X_2, ..., X_K]$.

Caso en que las "K" variables son Continuas

La Probabilidad del evento **"ε"** conjunto, debe definirse en intervalos dentro del rango **"R"**, de las variables componentes.

1. $f(x_1, x_2, ..., x_K) \geq 0$ para todo $[x_1, x_2, ..., x_K] \in R$
2. $f(x_1, x_2, ..., x_K) = 0$ para cualquier otro valor fuera de R
3. $\iint_R ..... \int f(x_1, x_2, ..., x_K) dx_K .....dx_2 dx_1 \equiv 1$
4. Los valores $([x_1, x_2, ..., x_K], f(x_1, x_2, ..., x_K))$ para todo el rango forman la **Función de Densidad Conjunta**, del vector original $[X_1, X_2, ..., X_K]$.
5. La **Probabilidad de un cierto intervalo dado**, digamos, $([a_1,b_1],[a_2,b_2],.....[a_K,b_K])$ para los componentes de la Función de Densidad Conjunta será la siguiente:

$$P(a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, ..., a_K \leq x_K \leq b_K) = \int_{a1}^{b1} \int_{a2}^{b2} ... \int_{ak}^{bk} f(x_1, x_2, ..., x_K) dx_K ... dx_2 dx_1$$

Distribuciones Conjuntas Marginales

Dentro de la Distribución Conjunta, podemos tener la necesidad de trabajar con la Distribución Individual de cada componente, para éstos casos definiremos la **Distribución Marginal**, sea **"x_k"** el componente el cual deseamos obtener su Distribución Marginal, sea $P(x_{ki})$ la probabilidad del elemento **"i"** del componente **"k"**, entonces la probabilidad $P(x_{ki})$ será la suma de todos los otros componentes con sus elementos.

$$P_k(x_{ki}) = \sum_{\substack{\text{todos los} \\ \text{componentes} \\ \text{excepto k, para} \\ \text{elemento i}}} P(x_1, x_2, ..., x_K)$$ para Distribuciones Discretas

🡐

Para el caso continuo, la **Distribución Marginal**, queda definida como la **Función de Densidad para la variable** **"x_k"** y será la solución de una integral múltiple sobre todos los componentes excepto **"k"**, con límites de **"- ∞ a ∞"** para cada uno.

$$f_k(x_k) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_K)\,dx_K \ldots dx_2 dx_1$$ excepto para componente k.

Para aterrizar un poco éstas generalidades, ejemplificaré la situación bivariada:

Caso Discreto

Consideremos una distribución discreta, con vector de componentes: $\vec{X} = [\vec{X}_1, \vec{X}_2]$

Con Distribución Conjunta siguiente: X1➔ Renglones, X2➔ Columnas, P(x1,x2)➔ Prob.Conj.

	0	1	2	3	4	P(X1)
0	.00010	.00470	.00420	.00065	.00035	**.01000**
1	.00900	.42300	.37800	.05850	.03150	**.90000**
2	.00060	.02820	.02520	.00390	.00210	**.06000**
3	.00030	.01410	.01260	.00195	.00105	**.03000**
P(X2)	**.01000**	**.47000**	**.42000**	**.06500**	**.03500**	**1.00000**

Tabla No. 4-1

Tenemos **2 distribuciones marginales**, éstas son:

Para X1 ➔ (0,.010), (1,.900), (2,.060), (3,.030)

Para X2 ➔ (0,.010), (1,.470), (2,.420), (3,.065), (4,.035)

Caso Continuo

Consideremos una densidad conjunta, con vector de componentes: $\vec{X} = [\vec{X}_1, \vec{X}_2]$

Cuya función de densidad es: $f(x_1,x_2) = 1/400$, para el intervalo: $x_1 \in [0,2]$ y $x_2 \in [0,200]$

Tenemos **2 distribuciones marginales**, éstas son:

Para X1 ➔ $$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\,dx_2 = \int_0^{200} \frac{dx_2}{400} = \frac{1}{2}$$

Para X2 ➔ $$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\,dx_1 = \int_0^{2} \frac{dx_1}{400} = \frac{1}{200}$$

Medias y Varianzas de Distribución Conjunta

Las Medias y Varianza en distribuciones conjuntas, se determinan a partir de sus distribuciones marginales, exactamente como en el caso univariado.

Obtención de la Distribución Individual de "Y" que es una Función de los componentes De la Distribución Conjunta Original

En ésta síntesis, se desarrollará sólo para el caso bivariado, tanto Discreto como Continuo.

Sea **"Y"** una función de X_1 y X_2: $Y = H(X_1,X_2)$.

Las variables X_1,X_2 dependen sin duda alguna del resultado del experimento original **"e"**, y **"Y"** es variable aleatoria con un rango de valores dependiente de los rangos de los componentes X´s.

Para el caso Discreto, puede obtenerse directamente de la Distribución Conjunta de las X´s, para ejemplificarlo, tomaré la distribución P(x1,x2) de la **tabla No. 4–1**, considerando que:

$Y = H(X_1,X_2) = 2X_1 + X_2$ entonces, los valores de **"y"** son: $y = 2x_1 + x_2$.

Los valores que puede tomar **"Y"** son **[0,10]** ya que es funcional a las X´s. La Distribución para la variable **"Y"**, la obtenemos a partir de los valores P(x1,x2).

y_i	(x_1,x_2)	Sumas $P(x_1,x_2)$	$P_y(y_i)$
0	(0,0)	.00010	.00010
1	(0,1)	.00470	.00470
2	(0,2);(1,0)	.00420+.00900	.01320
3	(0,3);(1,1)	.00065+.42300	.42365
4	(0,4);(1,2);(2,0)	.00035+.37800+.00060	.37895
5	(1,3);(2,1)	.05850+.02820	.08670
6	(1,4);(2,2);(3,0)	.03150+.02520+.00030	.05700
7	(2,3);(3,1)	.00390+.01410	.01800
8	(2,4);(3,2)	.00210+.01260	.01470
9	(3,3)	.00195	.00195
10	(3,4)	.00105	.00105

Para el caso Continuo, es un poco más complicado que una transformación de una sola variable, pero el procedimiento general consta de **8 pasos**, que se ejemplificarán a continuación para el caso bivariado. Por la utilización del **"Jacobiano"** que es el Determinante de una matriz cuadrada de Derivadas Parciales, necesitamos un número par de variables totales, es decir, si tenemos 2 componentes X´s, serán 4.

Procedimiento General para obtener la Función de Densidad $g_Y(y)$

1. Tener la función que relaciona $Y = H_1(X_1,X_2)$, como la densidad conjunta es de 2 componentes, necesitaremos 4 variables (x_1, x_2, y, z).
2. Introducir la 4° variable **"z"** en una función que relacione fácilmente con los componentes, ésta función puede ser por conveniencia para acompletar el Jacobiano o puede ser también de interés, definámosla como: $Z = H_2(X_1,X_2)$, sin perder de vista que deseamos poder resolver **"Y"** para X_1, X_2 en términos de Y y Z.
3. Encontrar las ecuaciones inversas de las funciones H´s, dando como resultado las funciones siguientes: $x_1 = G_1(y, z)$ y $x_2 = G_2(y, z)$.
4. Obtener las derivadas parciales correspondientes, doy sin probar que existen y son continuas, las siguientes parciales:
$$\frac{\partial x_1}{\partial y}, \frac{\partial x_1}{\partial z}, \frac{\partial x_2}{\partial y}, \frac{\partial x_2}{\partial z}$$
5. Obtener la función conjunta original, en términos de **"y"** y **"z"**.
$$f\,[G_1(y, z), G_2(y, z)]$$
6. Obtener el Jacobiano $J(y, z)$, a partir del paso (4), resolviendo el siguiente Determinante.
$$J(y,z) = \begin{vmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{vmatrix}$$
7. Obtener la densidad conjunta en términos de (y, z), mediante la multiplicación de los resultados del paso (5) y paso (6).
$$m(y, z) = f\,[G_1(y, z), G_2(y, z)] \bullet |J(y, z)|$$

←

8. Obtener la **Función de Densidad** $g_Y(y)$, mediante la Integral sobre **"z"**.

$$g_Y(y) = \int_{-\infty}^{\infty} m(y,z)dz$$

Ejemplo: Considere que tenemos un vector de componentes $[X_1, X_2]$, son independientes las variables X´s, con funciones de densidad marginales siguientes:

$$f_1(x_1) = 2e^{-2x_1} \quad \text{y} \quad f_2(x_2) = 2e^{-2x_2}, \; x_1 \text{ y } x_2 > 0.$$

Por lo tanto, al ser independientes, tenemos una función conjunta de la forma:

$$f(x_1, x_2) = 4e^{-2(x_1+x_2)}$$

1. Por la naturaleza del proceso en estudio, nos interesa encontrar la densidad de la siguiente variable **"Y"**, $Y = H_1(X_1,X_2) = X_1/X_2$, entonces $y = x_1/x_2$.
2. Por conveniencia definimos $Z = X_1 + X_2$, entonces $z = x_1 + x_2$, para la 4° variable.
3. Encontramos las funciones inversas, y son las siguientes:

$$x_1 = \frac{yz}{(1+y)} \quad \text{y} \quad x_2 = \frac{z}{(1+y)}$$

4. Obtenemos las derivadas parciales:

$$\frac{\partial x_1}{\partial y} = \frac{z}{(1+y)^2}, \quad \frac{\partial x_1}{\partial z} = \frac{y}{(1+y)} \quad \text{y} \quad \frac{\partial x_2}{\partial y} = -\frac{z}{(1+y)^2}, \quad \frac{\partial x_2}{\partial z} = \frac{1}{(1+y)}$$

5. Obtenemos la función conjunta original, en términos de **"y"** y **"z"**.

$$f\,[G_1(y, z), G_2(y, z)] = 4e^{-2z}$$

6. Obtenemos el Jacobiano $J(y,z)$.

$$J(y,z) = \begin{vmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{vmatrix} = \begin{vmatrix} \frac{z}{(1+y)^2} & \frac{y}{(1+y)} \\ -\frac{z}{(1+y)^2} & \frac{1}{(1+y)} \end{vmatrix} = \frac{z}{(1+y)^2}$$

7. Obtenemos la densidad conjunta en términos de (y, z), mediante la multiplicación de los resultados del paso (5) y paso (6).

$$m(y, z) = f\,[G_1(y, z), G_2(y, z)] \bullet |J(y, z)| = \left[\frac{z}{(1+y)^2}\right]\left[4e^{-2z}\right]$$

8. Obtenemos la **Función de Densidad** $g_Y(y)$, mediante la Integral sobre **"z"**.

$$g_Y(y) = \int_{-\infty}^{\infty} m(y,z)dz = \frac{1}{(1+y)^2} \qquad \text{para } y > 0.$$

←

Distribuciones Conjuntas Condicionales

En muchos casos de aplicación, puede ser importante encontrar la distribución que cumple un valor dado de los demás componentes conjuntos, se aplica el concepto de condicionalidad, en el caso bivariado las expresiones son en la siguiente forma:

Discreto: $P(X_2/x_{1(i)}) = \dfrac{P(x_{1(i)}, X_2)}{P(x_{1(i)})}$ ó $P(X_1/x_{2(j)}) = \dfrac{P(X_1, x_{2(j)})}{P(x_{2(j)})}$

Continuo: $f_2(X_2/x_1) = \dfrac{f(x_1, X_2)}{f_1(x_1)}$ ó $f_1(X_1/x_2) = \dfrac{f(X_1, x_2)}{f_2(x_2)}$

5. Distribuciones que se derivan de la Distribución Normal

Las siguientes distribuciones son derivadas de la Distribución original Normal, son útiles ya que nos permiten generar distribuciones de Muestreo, éstas son, Chi – Cuadrada (o también llamada Ji – Cuadrada), "t" de Student y "F" de Fisher.

Distribución Chi – Cuadrada

Sea **"x"** variable aleatoria Chi – Cuadrada, entonces su función de Densidad de Probabilidad es:

$$f_{\chi^2}(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\left(\frac{k}{2}-1\right)} e^{-\frac{x}{2}}$$ en donde **"k"** son los grados de libertad y **x > 0**.

= 0 para cualquier otro valor de **"x"**.

La **función de densidad chi – cuadrada**, es un caso particular de la Gamma cuando **"r = k/2"** y **"λ = ½"**.

A su vez, la **función de densidad chi – cuadrada**, puede verse como la suma de **"k"** variables normales estandarizadas ($\mu = 0$, $\sigma^2 = 1$), cada una elevada al cuadrado.

$\chi_k^2 = \sum_{i=1}^{k} Z_i^2$ en donde $Z_i \sim N(0,1)$.

Principio de Aditividad de la Chi – Cuadrada

Si tenemos **"p"** variables aleatorias independientes Chi – Cuadradas, con $k_1, k_2, \ldots, k_p$ grados de libertad respectivamente, y definimos una variable **"Y"** como la suma de las variables chi – cuadradas, entonces tenemos que **"Y"** sigue también una Chi – Cuadrada con **"k"** grados de libertad que son la suma de los grados individuales.

$Y = \chi_1^2 + \chi_2^2 + \ldots\ldots + \chi_p^2$ con grados de libertad siguientes: $k = \sum_{i=1}^{p} k_i$

Media, Varianza y FGM de la Chi – Cuadrada

Su media es: $\mu = k$

Su Varianza es: $\sigma^2 = 2k$

Su FGM es: $Mx(t) = \left(\dfrac{1}{1-2t}\right)^{\frac{k}{2}}$ para $t < ½$

🡐

Distribución "t" – Student

Otra distribución de considerable importancia práctica es la que se forma a partir de la razón (o división) de una variable normal estándar entre la raíz cuadrada de una Chi – Cuadrada independiente dividida entre sus grados de libertad, definida en la siguiente forma:

Sea variable **Normal estándar** **"Z" ~ N(0,1)** y sea **"V"** **variable chi – cuadrada**, en donde son independientes, entonces la nueva variable queda definida en la siguiente forma:

$$T = \frac{Z}{\sqrt{V/k}}$$ en donde **"T"** es variable Student, con **"k"** grados de libertad.

Su Función de Densidad se puede obtener a partir del procedimiento general visto en la sección 4.1 de Distribuciones Conjuntas, ya que **"Z"** y **"V"** son variables independientes, correspondientes en el procedimiento a X_1 y X_2.

La variable **"Y"** del procedimiento, corresponde a nuestra variable **"T"** y definimos una nueva variable **"U = V"**, que corresponde a la **"Z"** del procedimiento y obtenemos lo siguiente:

Su Densidad es: $$f(t) = \frac{\Gamma\left[\frac{(k+1)}{2}\right]}{\Gamma\left[\frac{k}{2}\right]} \cdot \frac{1}{\sqrt{k\pi}} \cdot \frac{1}{\left[\frac{t^2}{k}+1\right]^{\frac{(k+1)}{2}}} \qquad -\infty < t < \infty.$$

Su Media y Varianza son: $\mu = 0$ y $\sigma^2 = k/(k-2)$ para $k > 2$

Su comportamiento Gráfico es semejante a la Normal Estándar, centrada en **t = 0**, pero con grados de libertad finitos, tiene más probabilidad en los extremos, consecuentemente la ordenada en el origen es menor que en la Normal Estándar; sólo cuando sus grados de libertad tienden a infinito, la Normal Estándar es su forma límite.

Distribución "F" de Fisher

Ésta distribución tiene también buen interés práctico para la resolución de problemas, está definida como la razón (o división) de 2 variables chi – cuadradas Independientes, divididas entre sus respectivos grados de libertad. Si consideramos que **"U"** y **"V"** son distribuciones chi – cuadradas independientes, con **"m"** y **"n"** grados de libertad correspondientes, entonces partiendo de las distribuciones marginales, su función de densidad conjunta es:

$$f_U(u) = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} u^{\left(\frac{m}{2}-1\right)} e^{-\frac{u}{2}} \quad \text{y} \quad f_V(v) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} v^{\left(\frac{n}{2}-1\right)} e^{-\frac{v}{2}}$$

$$f(u,v) = \frac{1}{2^{\left(\frac{m+n}{2}\right)}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} u^{\left(\frac{m-2}{2}\right)} v^{\left(\frac{n-2}{2}\right)} e^{-\frac{1}{2}(u+v)}$$

Sea la transformación de la Dist. De Fisher: $$F = \frac{U/m}{V/n}$$

Su Función de Densidad se puede obtener a partir del procedimiento general visto en la sección 4.1 de Distribuciones Conjuntas, ya que **"U"** y **"V"** son variables independientes, correspondientes en el procedimiento a X_1 y X_2.

La variable **"Y"** del procedimiento, corresponde a nuestra variable **"F"** y definimos una nueva variable **"W = V"**, que corresponde a la **"Z"** del procedimiento y obtenemos lo siguiente:

🡐

$$f_F(f) = \frac{\Gamma\left[\frac{m+n}{2}\right]}{\Gamma\left[\frac{m}{2}\right]\Gamma\left[\frac{n}{2}\right]} \bullet \left(\frac{m}{n}\right)^{\frac{m}{2}} \bullet \frac{f^{\left(\frac{m-2}{2}\right)}}{\left[1+\left(\frac{m}{n}\right)f\right]^{\left(\frac{m+n}{2}\right)}} \qquad \text{para } f > 0.$$

Notar que el 1° término de los tres que se multiplican corresponde al recíproco de la función matemática Beta, vista en la sección 3.5 de ésta síntesis, con **a = m/2** y **b = n/2**.

El valor del Jacobiano correspondiente es: $J(f,w) = \frac{m}{n}v$

Su Media y Varianza son: $\mu = \frac{n}{n-2}$, con $n > 2$ y $\sigma^2 = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$, con $n > 4$.

Su comportamiento Gráfico es semejante a la chi – cuadrada es asimétrica hacia la derecha, pero su media se mantiene ligeramente mayor que **"1"**, al tener dos parámetros de grados de libertad (m,n) la **"F"** tiene mayor flexibilidad en su extremo derecho.

Dos Propiedades de la "F"

1. Si **"α"** corresponde al área a partir del valor de $F_{\alpha,m,n}$ hacia el infinito y deseamos obtener el valor de F que corresponde al complemento del área, se puede obtener ya que es el recíproco de "F" con los grados de libertad permutados, es decir[1]:

$$F_{(1-\alpha),n,m} = \frac{1}{F_{\alpha,m,n}}$$

2. Si **"X"** es una variable aleatoria con función de densidad de **"F"** de Fisher, con **"m"** y **"n"** grados de libertad en el numerador y denominador respectivamente; y se define una nueva variable **"W"** con la siguiente expresión:

$$W = \frac{\left(\frac{m}{n}\right)X}{1+\left(\frac{m}{n}\right)X}$$

La nueva variable **"W"** sigue una Distribución Beta con parámetros **a = m/2** y **b = n/2**.

Se puede demostrar ésta propiedad (2), utilizando la condición (3) de la sección 1.17 y substituyendo las expresiones.

En donde "X" tiene la siguiente expresión: $X = \frac{n}{m} \cdot \frac{W}{1-W}$, y usando la condición (3), planteada para éste caso es: $f_W(w) = f_X(X)\frac{dx}{dw}$

[1] La demostración se encuentra en el Graybill and Boes, Págs. 248 a 249.

6. Anexos a los Temas (1, 2, 3, 4 y 5)

Concepto de Independencia en las Variables Aleatorias

En ésta sección, tocaré las nociones de Independencia desde el punto de vista del Proceso Estocástico, que en forma general, Proceso Estocástico es todo aquel que el resultado del estado actual, depende de los resultados del estado anterior.

Las nociones de Independencia, y de variables aleatorias independientes, son conceptos estadísticos muy útiles e importantes, en la sección 1.7 de ésta síntesis se introdujo el concepto de sucesos o eventos independientes, ahora estamos interesados en la definición de **variables aleatorias independientes**.

Cuando el resultado de una variable, digamos, $\mathbf{X_1}$, no modifica el resultado de $\mathbf{X_2}$ y viceversa, se afirma que las variables $\mathbf{X_1}$ y $\mathbf{X_2}$ son independientes.

Definiciones

1. Si $[\mathbf{X_1, X_2, ..., X_K}]$ es un vector de variables aleatorio discreto K-dimensional, planteamos que las X´s son **Independientes**, **si solo si**, se cumple la siguiente expresión:

$$p(x_1, x_2, ..., x_K) = \prod_{i=1}^{K} p_i(x_i)$$

2. Si $[\mathbf{X_1, X_2, ..., X_K}]$ es un vector de variables aleatorio continuo K-dimensional, planteamos que las X´s son **Independientes**, **si solo si**, se cumple la siguiente expresión:

$$f(x_1, x_2, ..., x_K) = \prod_{i=1}^{K} f_i(x_i)$$

Utilizando los conceptos de Independencia y Condicionalidad, para el caso bivariado tenemos:

1. **Opción Discreta:**

$$P(X_{2(j)}/x_{1(i)}) = \frac{P(x_{1(i)}, X_{2(j)})}{P_1(x_{1(i)})} = p(x_{2(j)}) \quad \text{ó} \quad P(X_{1(i)}/x_{2(j)}) = \frac{P(X_{1(i)}, x_{2(j)})}{P_2(x_{2(j)})} = p(x_{1(i)})$$

2. **Opción Continua:**

$$f_2(X_2/x_1) = \frac{f(x_1, X_2)}{f_1(x_1)} = f_2(x_2) \quad \text{ó} \quad f_1(X_1/x_2) = \frac{f(X_1, x_2)}{f_2(x_2)} = f_1(x_1)$$

Si son independientes entre si la X´s, encontraremos los valores de las distribuciones marginales, en el caso bivariado.

Para el caso K-dimensional y siendo independientes entre si los componentes X´s, al buscar la condicionalidad, digamos bajo **"$\mathbf{X_1}$"**, encontramos la función conjunta de los demás componentes, dado el valor utilizado de **"$\mathbf{x_1}$"**.

Como ejemplo, manejaré el vector de 3 componentes $[\mathbf{X_1, X_2, X_3}]$, continuo.

$$f((X_1, X_2, X_3)/x_1) = \frac{f(x_1, X_2, X_3)}{f_1(x_1)} = f(x_2, x_3)$$ dado para ese valor de **"$\mathbf{x_1}$"** en particular.

Concepto de Covarianza y Correlación

Si tenemos dos variables aleatorias **"X"** y **"Y"**, la covarianza es una medida de la asociación lineal que pueda existir entre dichas variables. El coeficiente de correlación tiene involucrada a la covarianza y describe el grado de asociación lineal que existe entre ambas variables, éste es adimensional y las expresiones poblacionales de ambos son las siguientes:

$\leftarrow$

$$\mathbf{Cov(X,Y) = \sigma_{x\,y} = E[(X - E(X))(Y - E(Y))] = E(X \bullet Y) - [E(X)E(Y)]}$$ **Covarianza x,y**

$$\rho = \frac{Cov(X,Y)}{\sqrt{V(X)} \cdot \sqrt{V(Y)}} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$ **Coeficiente de correlación**

Si tenemos un tamaño de muestra (**n**), la expresión muestral para el coeficiente de correlación es la siguiente:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{X})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{Y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}}$$

Si **[X_1, X_2]** es un vector de variables aleatorias continuas **independientes**, entonces su Covarianza vale cero: $\mathbf{Cov(X_1,X_2) = 0}$ y el coeficiente de correlación entre ambas vale cero: $\rho = \mathbf{0}$.
Lo inverso a este enunciado **no necesariamente es cierto**, en ese caso sólo se puede afirmar que las variables **no están correlacionadas**.
Al ser independientes las variables X´s, la densidad conjunta de ambas es igual al producto de sus densidades marginales, entonces tenemos:

$$E(X_1 \cdot X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1,x_2)dx_2 dx_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f_1(x_1)dx_1 \cdot x_2 f_2(x_2)dx_2 =$$
$$= \left[\int_{-\infty}^{\infty} x_1 f_1(x_1)dx_1\right] \cdot \left[\int_{-\infty}^{\infty} x_2 f_2(x_2)dx_2\right] = E(X_1) \cdot E(X_2)$$

Los valores de la Covarianza y del Coeficiente de Correlación son:

$$\mathbf{Cov(X_1,X_2) = \sigma_{1\,2} = E[(X_1 - E(X_1))(X_2 - E(X_2))] = E(X_1 \cdot X_2) - [E(X_1)E(X_2)] = 0}$$
$$\rho = \mathbf{0}$$

El coeficiente de correlación, tiene únicamente valor en el intervalo cerrado **[–1, +1]**, es decir:
$\mathbf{-1 \le \rho \le +1}$.
Para lograr la comprobación de los valores que puede tomar el coeficiente de correlación, necesitamos recurrir a la **"Desigualdad de Cauchy – Schwart"**[2], al concepto matemático del Discriminante[3] y al planteamiento de una función cuadrática de los valores esperados[4].

La desigualdad de **Cauchy – Schwart**, plantea que: $\mathbf{(E(X_1X_2))^2 \le E(X_1^2) \cdot E(X_2^2)}$ con el valor de igualdad **si y solo si** la relación de las variables $\mathbf{X_1}$ y $\mathbf{X_2}$ es matematicamente lineal, con probabilidad de **"1"**, en la relacion $\mathbf{X_1 = dX_2}$, siendo **"d ≠ 0"** una constante, es decir:
$\mathbf{P(X_1 = dX_2) = 1}$.

Si definimos una función: $\mathbf{Q(t) = E\left(\left[(X_1 - \mu_1) - t(X_2 - \mu_2)\right]^2\right)}$, notar que ésta función solo puede tomar valores no negativos $\mathbf{Q(t) \ge 0}$, al desarrollarla obtenemos:

$\leftarrow$

[2] **"Introduction to the Theory of Statistics"**, Mood, Graybill and Boes, Pág. 162.
[3] **"Calculus and Analytic Geometry"**, Thomas, páginas 352 a 353.
[4] **"Probabilidad y Estadística"**, Montgomery, páginas 148 a 149.

$$Q(t) = E[(X_1 - \mu_1)^2] - 2tE[(X_1 - \mu_1)(X_2 - \mu_2)] + t^2E[(X_2 - \mu_2)^2]$$

Una relación cuadrática en función de **"t"**, con la siguiente expresión:

$$Q(t) = at^2 + bt + c \geq 0$$

En donde las constantes de la relación son:

$$a = E[(X_2 - \mu_2)^2] = V(X_2)$$

$$b = -2E[(X_1 - \mu_1)(X_2 - \mu_2)] = -2Cov(X_1, X_2)$$

$$c = E[(X_1 - \mu_1)^2] = V(X_1)$$

El discriminante, resultado de la función Q(t), aunado a la aplicación de la desigualdad de Cauchy – Schwart, sólo puede tener valores menores o igual a cero, es decir:

$$Dis = b^2 - 4ac = 4[Cov(X_1, X_2)]^2 - 4V(X_1)V(X_2) \leq 0$$

Para el caso en que el discriminante vale cero, podemos encontrar el valor **"t_0"** que corresponde a la constante **"d"** descrita en la desigualdad de Cauchy – Schwart, (con probabilidad **"1"**).

Para el caso en que el discriminante es menor que cero, no tiene solución real, pero si en el sistema complejo basado en $i = \sqrt{-1}$.

Retornando a la expresión del discriminante, reagrupando los términos de Covarianza y Varianzas, tenemos:

$$\rho^2 = \frac{[Cov(X_1, X_2)]^2}{V(X_1)V(X_2)} \leq 1 \quad \rightarrow \quad -1 \leq \rho \leq +1$$