VII. VARIABLES ALEATORIAS DISTRIBUIDAS CONJUNTAMENTE.

7.1 Definiciones

Caso de V.A. Discretas:

Para cualquier par de V.A. ***X e Y***, la función de distribución acumulada conjunta de *X e Y* es definida como:

$$F(a,b) = P\{ X \text{ £ } a, Y \text{ £ } b \} \text{ , para todo } -\infty < a,b < +\infty$$

La función de distribución de *X* puede ser obtenida desde la función de distribución conjunta de *X e Y* de la siguiente manera:

$$F_X(a) = P(X \text{ £ } a) = P\{ X \text{ £ } a, Y \text{ £ } +\text{¥} \} = F(a,\text{¥})$$

De manera similar se obtiene $F_Y(b) = F(\text{¥},b)$.

La función de densidad de la conjunta viene dada por:

$$p(x,y) = P\{ X = x, \ Y = y \}$$

con la cual puede ser obtenida la f.d.p. de *X e Y* de la manera siguiente:

$$p_X(x) = \sum_{y/p((x,y)>0} p(x,y) \quad \text{y} \quad p_Y(y) = \text{å}_{x/p((x,y)>0} \ p(x,y)$$

Caso de V.A. Continuas:

Definición 7.1

X e Y son V.A. continuas conjuntamente distribuidas si $ ***f(x,y)***, conocida como función de densidad conjunta de *X e Y*, definida para todo $x,y \in$ R y que cumple con, $\forall$A,B $\subseteq$ R,

$$P\{ X \ \text{Î} \ A, Y \ \text{Î} \ B \} = \int_A \int_B f(x,y) \, dx \, dy.$$

Con la función de densidad conjunta podemos obtener la f.d.p. de X de la manera siguiente:

$$P\{\ X\ Î\ A\ \} =\ P\{\ X\ Î\ A,\ Y\ Î\ (-¥,+¥\ )\ \} = \iint_A f(x,y)\,dx\,dy = \int_A f_X(x)\,dx$$

Donde $f_X(x) = \int_{-\infty} f(x,y)\,dy$ de la misma forma tenemos $f_Y(y) = \int_{-\infty} f(x,y)\,dx$.

<u>Esperanza matemática de una función g(x,y) de dos V.A. X e Y.</u>

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)p(x,y) \quad : para\ el\ caso\ discreto$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy \quad : para\ el\ caso\ continuo$$

$$E[X+Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy +$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy = \int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f(x,y)dy\right)dx + \int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} f(x,y)dx\right)dy$$

$$= \int_{-\infty}^{\infty} x\ f_X(x)dx + \int_{-\infty}^{\infty} f_Y(y)dy = E[X] + E[Y]$$

e1) Si $g(x,y) = x + y$, entonces:

Igualmente se trata el caso discreto.

Este último resultado, combinado con la propiedad $E[aX] = aE[X]$ produce:

$$E[aX+bY] = aE[X] + bE[Y]$$

En el caso de tener distribuciones conjuntas definidas para n V.A., digamos $X_1,X_2,...,X_n$, tendremos que para todo $a_1,a_2,..a_n$ se cumple:

$$E[a_1X_1 + a_2X_2 +..+ a_nX_n] = a_1E[X_1] + a_2E[X_2] + .. + a_nE[X_n]$$

7.2 Variables Aleatorias Independientes

Definición 7.2

X e Y son independientes si para todo a,b tenemos:

$$P\{X £ a, Y £ b\} = P\{X £ a\} * P\{Y £ b\}.$$

Luego podemos decir que *X e Y* son independientes sii los eventos $E_a=\{x/ x£a\}$ y $F_b=\{y/ y£b\}$ son independientes, para todo *a, b*..

En términos de la función de densidad conjunta podemos decir que:

X e Y son independientes sii $F(a,b) =F_X(a) * F_Y(b)$, para todo *a,b*.

Entonces,

para el caso discreto, $p(x,y) = p_X(x) * p_Y(y)$ y

para el caso continuo, $f(x,y) = f_X(x) * f_Y(x)$.

Lema.
Si *X e Y* son independientes, entonces para cualquier par de funciones *h(.)* y *g(.),* en *X* y en *Y* respectivamente tendremos: $E[g(X)h(Y)] = E[g(X)]*E[h(Y)]$.

Demostración.
Supongamos que *X e Y* son continuas y que existe una distribución conjunta *f(x,y)* para ellas, entonces:

$$E[g(X)h(Y)] = \iint g(x)h(y)f(x,y)dxdy =$$

(por independencia) = $\iint g(x)h(y)f_X(x)f_Y(y)dxdy = \int h(y)f_Y(y)dy * \int g(x)f_X(x)dx$

$$= E[h(Y)]*E[g(X)].$$

(El caso discreto se muestra de manera similar).

Calculo de $F_{X+Y}(a)$ cuando F_X y F_Y son conocidas y *X e Y* son independientes

(F_{X+Y} se le conoce como **Convolución** de F_X y F_Y).

Sean $f_X(.)$ y $f_Y(.)$ f.d.p de *X e Y* respectivamente. Luego,

$$F_{X+Y}(a) = P(X+Y \leq a) = \iint_{X+Y\leq a} f(x,y)dxdy$$

$$\overset{indep}{=} \iint_{X+Y\leq a} f_X(x) f_Y(y)dxdy$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{a-y} f_X(x)dx\right] f_Y(y)dy = \int_{-\infty}^{\infty} F_X(a-y) f_Y(y)dy$$

$$f_{X+Y}(a) = \frac{d}{da}\int_{-\infty}^{\infty} F_X(a-y)\, f_Y(y)dy$$

$$= \int_{-\infty}^{\infty} \frac{dF_X(a-y)}{da} f_Y(y)dy = \int_{-\infty}^{\infty} f_X(a-y)\, f_Y(y)dy$$

A continuación obtendremos la f.d.p. de $X+Y$ diferenciando $F_{X+Y}(a)$.

<u>Ejemplo 1.</u>

Sean ***X y Y*** V.As. independientes uniformemente distribuidas en el intervalo (0,1). Encontrar la función de densidad f_{X+Y}.

<u>Solución</u>.

$$f_X(a) = f_Y(a) = \begin{cases} 1 \text{ si } 0 < a < 1 \ \ y \\ 0 \text{ en caso contrario} \end{cases}$$

$$\Rightarrow f_{X+Y}(a) = \int_0^1 f_X(a-y)dy$$

Analizando por intervalos donde podría estar “a”:

para $0 \leq a \leq 1$ tenemos que:

$$f_{X+Y}(a) = \int_0^a dy = a$$

y para $1 < a \leq 2$ tenemos que:

$$f_{X+Y}(a) = \int_{a-1}^{1} dy = 2 - a$$

$$\_> \quad f_{X+Y}(a) = \begin{cases} a & 0 \leq a \leq 1 \\ 2 - a & 1 < a < 2 \\ o & \text{en caso contrario} \end{cases}$$

Ejemplo 2.

Sean X *e* Y V.A.s del tipo Poisson independientes con parámetros λ_1 y λ_2 respectivamente. Encuentre $f_{X+Y}(a)$.

Solución.

$$P(X+Y=n)=\sum_{k=0}^{n} P(X=k, Y=n-k)$$

$$\text{ya que } \{(X,Y) / X+Y=n\}=\bigcup_{k=0}^{n}\{(X,Y) / X=k, Y=n-k\}$$

Entonces:

$$f_{X+Y}(n)=\sum_{k=0}^{n} P(X=k)P(Y=n-k)$$

$$=\sum_{k=0}^{n}\left[e^{-\lambda_1}\frac{\lambda_1^{k}}{k!}\right]\left[e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!}\right]=e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{n}\frac{\lambda_1^{k}\lambda_2^{n-k}}{k!(n-k)!}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^{k}\lambda_2^{n-k}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\binom{n}{k}\lambda_1^{k}\lambda_2^{n-k}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^{n}$$

lo que representa una Poisson con parámetro $\lambda=\lambda_1+\lambda_2$.

7.3 Distribuciones de V.A Condicionadas.

Caso discreto:

La función de densidad condicional de una V.A. discreta viene dada por:

$$p_{x/y}(x/y) = \begin{cases} P\{X=x/Y=y\} = P\{X=x,Y=y\} / p\{Y=y\} = p(x,y)/p_Y(y), & \text{para } p_Y(y) > 0 \\ 0 & \text{si no} \end{cases}$$

Asimismo la función acumulada condicional viene dada por:

$$F_{X/Y}(x/y) = P\{X£x \,/\, Y=y\} = å_{a£x}\, p_{x/y}(a/y).$$

<u>Ejemplo 1.</u>
Sean X e Y V.As. con $p(x,y)$ definidas por:

$p(0,0) = 0.4,\ \ p(0,1) = 0.2,\ \ p(1,0) = 0.1\ \ y\ \ p(1,1) = 0.3$.

Calcule $P(X = x / Y = 1)$.

<u>Solución.</u>

$P(X = x / Y = 1) = P(x,1) / P_Y(1)$. Ahora calculamos,

$P_Y(1) = å_x\, p(x,1) = p(0,1) + p(1,1) = 0.5$, por lo tanto:

$p_{x/y}(0/1) = p(0,1) / 0.5 = 2/5 \quad y \ \ p_{x/y}(1/1) = p(1,1) / 0.5 = 3.5$

<u>Ejemplo 2.</u>
Sean X *e* Y V.As. que siguen distribuciones Poisson independientes con parámetros l_1 *y* l_2 respectivamente. Encuentre la función de distribución condicional de X dado que $X + Y = n$

<u>Solución:</u>

$P\{X=k/X+Y=n\} = P\{X=k,X+Y=n\} / P\{X+Y=n\} = P\{X=k,Y=n-k\} / P\{X+Y=n\}$
y por independencia es

$$= (P\{X=k\}\, P\{Y=n-k\}) / P\{X+Y=n\}.$$

Como $P\{X+Y=n\}$ es conocida ya que fue encontrada en un ejercicio de la sección anterior, tendremos:

$$P\{X=k/X+Y=n\} = \frac{e^{-\lambda_1}\lambda_1^k}{k!}\,\frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}\left[\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}\right]^{-1}$$

$$\frac{n!}{(n-k)!k!}\,\frac{\lambda_1^k\,\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n} = \binom{n}{k}\left[\frac{\lambda_1}{\lambda_1+\lambda_2}\right]^k\left[\frac{\lambda_2}{\lambda_1+\lambda_2}\right]^{n-k}$$

Lo cual corresponde a una función de densidad de una distribución Binomial con parámetros n y

$$p = \lambda_1/(\lambda_1 + \lambda_2)$$

Caso continuo:

La función de densidad de una condicional es dada por:

$$f_{x/y}(x/y) = P\{x \le X \le x+dx \,/\, y \le Y \le y+dy\}$$

$$= f(x,y) \,/\, f_Y(y), \quad para\ f_Y(y) > 0$$

Asimismo la función acumulada condicional viene dada por:

$$P\{ X \in A \,/\, Y=y \} = \int_A f_{X/Y}(x/y)dx$$

En particular con A=(-∞,a] tendremos:

$$F_{X/Y}(a/y) = P\{X \le a \,/\, Y=y\} = \int_{-\infty}^{a} f_{x/y}(x/y)dx.$$

Ejemplo 1.

Calcule,

$$f_{X/Y}(x/y) \quad \text{dado} \quad f(x,y) = \begin{cases} x(2-x-y) & 0<x<1,\ 0<y<1 \\ 0 & \text{en caso contrario} \end{cases}$$

Solución.

$$f_{X/Y}(x/y) = \frac{f(x,y)}{f_Y(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx}$$

$$= \frac{x(2-x-y)}{\int_0^1 x(2-x-y)dx} = \frac{x(2-x-y)}{\frac{2}{3}-\frac{y}{2}}$$

$$= \frac{6x(2-x-y)}{4-3y} \quad \text{con} \quad 0<x,\ y<1$$

Ejemplo 2.

$$P\{X > 1 / Y = y\} \;\; con \;\; f(x,y) = \begin{cases} \dfrac{e^{-\frac{x}{y}} e^{-y}}{y} & 0 < x,y < \infty \\ 0 & caso\ contrario \end{cases}$$

Calcule,

<u>Solución.</u>

$$P\{X > 1/Y = y\} = \int_1^{\infty} f_{X/y}(x/y)dx$$

$$f_{x,y}(x/y) = \frac{f(x,y)}{f_Y(y)} = \frac{\dfrac{e^{-\frac{x}{y}} e^{-y}}{y}}{e^{-y}\displaystyle\int_0^{\infty} \frac{1}{y} e^{-\frac{x}{y}} dx} = \frac{1}{y} e^{-\frac{x}{y}}$$

entonces,

$$P\{X > 1/Y = y\} = \int_1^{\infty} \frac{1}{y} e^{-\frac{x}{y}} dx = -e^{-\frac{x}{y}} \Big|_1^{\infty} = e^{-\frac{1}{y}}$$

<u>7.4 Covarianza, correlación y regresión</u>

7.4.1 Covarianza

La covarianza entre dos V.A. se denota con *Cov(X,Y)* y se define de la manera siguiente:

$$Cov(X,Y) = E[(X-E(X))*(Y-E(Y))].$$

Nota:
Con este valor se mide en cierta forma el grado de dependencia que hay entre las dos variables.

Lema 7.
$Cov(X,Y) = E[XY] - E[X]*E[Y]$

Demostración.
$Cov(X,Y) = E[(X-E(X))*(Y-E(Y))] = E[XY - XE[Y] - YE[X] + E[X]E[Y]]$

$= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] = E[XY] - E[X]E[Y]$

Nota

Si *X e Y* son independientes, entonces por el lema anterior, $Cov(X,Y) = 0$.

Ejemplo.

Sean X e Y dos V.A. del tipo Bernoulli, definidas de la manera siguiente:

X = 1 si A sucede y 0 en caso contrario y
Y = 1 si B sucede y 0 en caso contrario.

Entonces, como la $Cov(X,Y) = E[XY]-E[X]E[Y]$ y *XY* es "1" solo si ambos simultáneamente son "1" tendremos:

$Cov(XY) = p(X=1,Y=1) - p_X(X=1)p_Y(Y=1)$.

Luego, podemos concluir que $Cov(X,Y) > 0$ sii $p(X=1,Y=1) > p_X(X=1)p_Y(Y=1)$, es decir:

$p(X=1,Y=1)/p_X(X=1) > p_Y(Y=1)$ ó $p(Y=1 / X=1) > p(y=1)$.

En otras palabras, podemos decir que la covarianza es positiva si el resultado $Y=1$ se hace más probable sí $X=1$ y viceversa.

VARIANZA.DE LA SUMA DE V.A.

A continuación presentaremos una expresión útil para la Varianza de la suma de 2 V.A. en función de la Covarianza.

$$VAR(X+Y) =$$

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

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

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

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

$$VAR(X) + VAR(Y) + 2Cov(X,Y).$$

Nota:
Obsérvese que si *X y Y* son independientes se tendremos que:

$$VAR(X+Y)=VAR(X)+VAR(Y).$$

Ejemplo:(Varianza de una V.A. Binomial)

Sean X_i (i=1,..n) V.A. del tipo Bernoulli independientes

y sea $X = X_1+X_2+..+X_n$ entonces:

$VAR(X) = \sum_n VAR(X_i)$ y como $VAR(X_i) = E(X_i^2) - (E(X_i))^2 = E(X_i) - (E(X_i))^2 = p - p^2$, entonces,

$VAR(X) = np(1-p).$

Caso generalizado.

$VAR(\sum_n X_i) = \sum_n VAR(X_i) + \sum\sum_{i<j} Cov(X_i X_j)$

7.4.2 CORRELACIÓN

Como se habrá notado la definición de Covarianza entre 2 V.As. depende de la unidad en que son medidas y en consecuencia la Covarianza de dos V.A. en metros y en kilómetros varía demasiado y por ello es difícil usarlas como medida de dependencia. Igualmente pasa si intentamos comparar la dependencia de una V.A. X respecto a otras dos V.A.s Y y Z. Para evitar este inconveniente aparece un nuevo concepto independiente de la unidad de medida que nos permite disponer de un indicador del grado de dependencia entre dos variables.

<u>Definición</u>

Se llama Coeficiente de Correlación entre dos V.A.s, al cociente:

$$r = \frac{Cov(X,Y)}{s(X)s(Y)}, \quad con\ s(X), s(Y) \neq 0$$

A continuación se va a mostrar que ρ está en el intervalo [-1,1].

Considerando,

$$VAR(aX + bY) =$$

$$E[(a(X - E[X]) + b(Y - E[Y]))^2] = a^2 VAR[X] + b^2 VAR[Y] + 2ab Cov(X,Y) \geq 0$$

Por ser la esperanza matemática de un cuadrado, esta expresión es siempre positiva y por lo tanto, con:

$$a = \frac{1}{s(X)}, \quad b = \pm \frac{1}{s(y)}$$

luego,

$$\frac{VAR[X]}{s^2(X)} + \frac{VAR[Y]}{s^2(Y)} \pm \frac{2Cov(X,Y)}{s(X)s(Y)} \geq 0$$

de donde:

a) $-1 \leq \rho \leq 1$

b) Si $\rho = \pm 1$, existen valores $a = a_0$, $b = b_0$ (no ambos nulos), para los cuales

$E[(a_0(X-E[X]) + b_0(Y-E[Y]))^2] = 0$, lo que exige que:

$$a_0(X-E[X]) + b_0(Y-E[Y]) = 0$$

Esta igualdad significa que los únicos pares de valores x_i, y_j de las variables X,Y que tienen probabilidad distinta de cero de ocurrir, son los que verifican $a_0(x_i-E[X]) + b_0(y_j-E[Y]) = 0$, es decir, los pares (x_i, y_j) son coordenadas de puntos pertenecientes a la recta de la ecuación anterior. Esto implica que entre X e Y hay una correspondencia funcional lineal, lo cual se puede formular en el siguiente lema.

Lema

El coeficiente de correlación es un numero real ρ comprendido entre -1 y +1, tal que $\rho = \pm 1$ implica que entre X e Y hay una dependencia funcional lineal.

Nota: Si X e Y son independientes, hemos visto que Cov(X,Y) = 0, entonces $\rho = 0$. Sin embargo esta condición necesaria no es suficiente para la independencia de las variables X e Y. Esto ultimo será mostrado en el siguiente ejemplo.

Ejemplo. Sea *U* la V.A. del número de un dado lanzado al azar, y sea *V* la variable análoga para un segundo dado, independiente del primero.
Consideremos la variable $X = U+V$ e $Y = U-V$.
Tendremos que: $E[X.Y] = E[U^2-V^2] = E[U^2] - E[V^2] = 0$. Por otra parte, también es $E[Y]=E[U]-E[V]=0$ Por consiguiente, la $Cov(X,Y) = 0$, de donde $r = 0$. Sin embargo, *X e Y* no son independientes, pues ellas toman valores pares o impares a la vez.

7.4.3 REGRESION (Extra Cursus)

Sean X e Y dos V.As., entonces para cada valor de X, por ejemplo $X=x_i$, la esperanza de Y, como hemos visto anteriormente, es un cierto valor:

$$y_i = E[Y / X = x_i] = \sum_j y_j \, f_{X/Y}(y_j / x_i)$$

$$f_{X/Y}(y_j / x_i) = \frac{f(x_i, y_j)}{f_X(x_i)}$$

donde:
Se tiene así una función y' cuyo dominio es el conjunto de valores de X, se llama la **regresión de Y sobre X**.

Si se trata de aproximar esta función mediante una expresión lineal y=ax+b, se utiliza para ello el método de mínimos cuadrados, que consiste en determinar las constantes a y b de manera que la esperanza de la diferencia de los valores de y' y la recta ax+b, es decir:

$$E[] = \sum_i \Big(\sum_j y_i\, f_{X/Y}(y_j / x_i) - ax_i - b \Big)^2 f_X(x_i)$$

sea mínima. Para determinar los valores de a y b que hacen mínimo E[], deben ser nulas las derivadas parciales de esta expresión respecto de a y b, y teniendo en cuenta las ecuaciones de que la suma de las probabilidades f(xi,yj) sobre todo su dominio es igual a 1, se puede concluir que:

$$E[Y] - b - aE[X] = 0, \quad E[X.Y] - bE[X] - aE[X^2] = 0$$

donde, aplicando la expresión de la Cov(X,Y) en función de las esperanzas de X, Y y X.Y se obtiene finalmente que:

$$a = Cov(X,Y) / VAR[X] \quad y \quad b = E[Y] - (Cov(X,Y) / VAR[X]) * E[X]$$

Estos son los coeficientes de la llamada **recta de regresión de Y sobre X**. El coeficiente a se le conoce como el **coeficiente de regresión**.