# III. VARIABLES ALEATORIAS.

## 3.1 Introducción

Frecuentemente en los experimentos el interés está en una función del resultado del experimento y no en el resultado propiamente dicho. Por ejemplo, en el lanzamiento de dos dados el interés esta en que la suma de los resultados sea igual a 7 y se es indiferente si este resulta de (1,6), (2,5), (3,4), (4,3), (5,2) o (6,1).
En otra situación, por ejemplo, en el contexto de comunicación a través de mensajes, el interés del analista podría centrarse en el numero de transmisiones exitosas de mensajes en un tiempo dado, mas que en el estado de la transmisión de cada mensaje.
Las cantidades de interés y mas formalmente las funciones de valores del espacio muestral: X : ω -----> R, son conocidas como **Variables Aleatorias (VA)**. Por ello, el valor de la variable aleatoria es determinado por el resultado del experimento y a esos valores se les asignará una probabilidad, como se muestra con los siguientes ejemplos.

**Ejemplos**

**e1)** Si X denota una VA definida como la suma de los resultados en el lanzamiento de 2 dados, entonces:

X : Ω -----> {2,...,12}

$$1 = P\{\bigcup_{n=2}^{12}\{X = n\}\} = \sum_{n=2}^{12} P(X = n)$$

donde:

P(X=2) = P{(1,1)} = 1/36
P(X=3) = P{(1,2),(2,1)} = 2/36
P(X=4) = P{(1,3),(2,2),(3,2)} = 3/36
P(X=5) = P{(1,4),(2,3),(3,2),(4,1)} = 4/36
. .
. .
P(X=12) = P{(6,6)} = 1/36

**e2)** X una VA que representa el número de transmisiones con éxito de mensajes enviados en un intervalo de tiempo dado, entonces:

X : $\Omega_t$ -----> **N**, donde $\Omega_t$: conjunto de mensajes transmitidos exitosamente en un intervalo de tiempo t.

**e3)** Sea Y la VA que representa el número de caras que aparecen luego de lanzar 2 monedas. Luego la VA Y puede tomar valores 0, 1 o 2, y sus valores probabílisticos son:

P(Y=0) = P{(s,s)} = 1/4
P(Y=1) = P{(s,c),(c,s)} = 2/4
P(Y=2) = P{(c,c)} = 1/4
------
Σ = 1

**e4)** Suponiendo que una moneda esta "cargada" y que la probabilidad de que salga cara es 'p', definamos una VA N de la siguiente manera:

N : Número de lanzamientos necesarios para que salga cara por primera vez.

Entonces:

P(N=1) = P{(c)} = p

$P(N=2) = P\{(s,c)\} = (1-p)p$
$P(N=3) = P\{(s,s,c)\} = (1-p)^2p$
. .
. .
. .
$P(N=n) = P\{.(s,s,...,s, c)\} = (1-p)^{n-1}p$

Lo cual es una función probabílistica, ya que es positiva y

$$P\{\bigcup_{n=1}^{\infty}\{N = n\}\} = \sum_{n=1}^{\infty} P(N = n) = p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\sum_{t=0}^{\infty}(1-p)^{t} = \frac{p}{1-(1-p)} = 1$$

Todos los ejemplos de VA hasta ahora ilustrados son **VA discretas.**
Por otro lado una **VA se dice continua** si toma valores de un conjunto continuo de posibles valores. Un ejemplo de VA continua es el tiempo de transmisión de un mensaje, la vida útil de un motor, asumiendo que la vida útil del motor toma un valor en un intervalo (a, b) de los reales.

**Definición**
Sea F(.), una función definida para cualquier valor real b de una VA X, con b en el intervalo $(-\infty, \infty)$ de la siguiente manera: $\mathbf{F(b) := P\{X \leq b\}}$, es conocida como **Función de distribución acumulada de X** y denota la probabilidad de que X pueda tomar un valor menor o igual a b.

**Propiedades de F(.)**
(i) F(b) es una función no decreciente en b
(ii) $\lim_{b \to \infty} F(b) = 1$
(iii) $\lim_{b \to -\infty} F(b) = 0$
(iv) F es estrictamente continua. Si $x_1, x_2, ...$ es una serie decreciente con limite en x, entonces

$$\lim_{n \to \infty} F(x_n) = F(x)$$

**Observaciones sobre F(.):**
* Si $a \leq b$ entonces el evento $\{x / x \leq a\}$ esta en $\{x / x \leq b\}$ y por lo tanto $F(a) \leq F(b)$ (propiedad i).

* Probabilidades sobre la VA X intervalos pueden ser deducidas en función de F(.).
Por ejemplo: Prob.$\{a < X \leq b\} = F(b) - F(a)$, para todo $a < b$; ya que $(-\infty, b] = (-\infty,a] \cup (a,b]$.

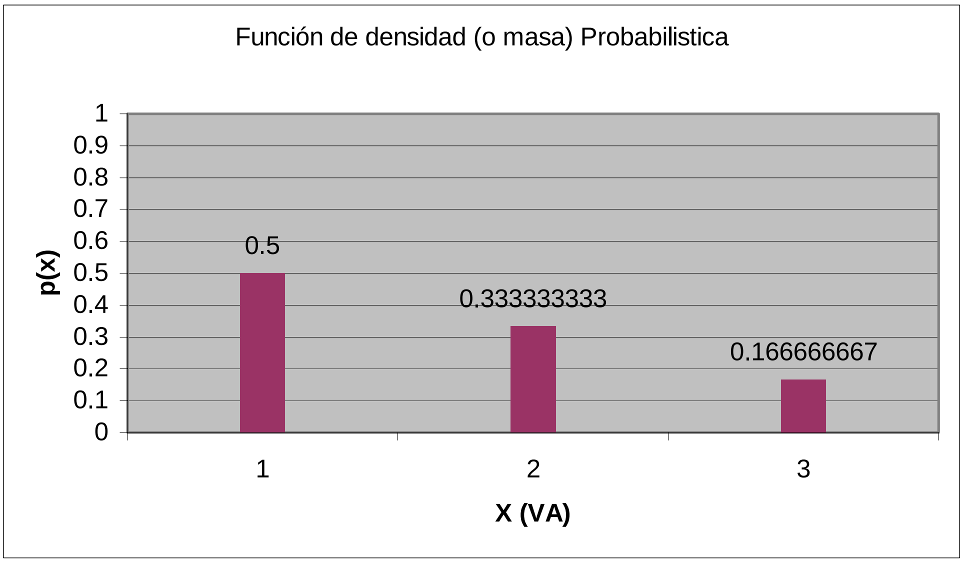
**Para VA discretas** definimos la función de densidad de X con $p(a) = P\{X=a\}$.
Si $X = \{x_1, x_2,... \}$ en Z, entonces $p(x_i) > 0$ con i=1,2,...
y $p(x) = 0$ para todo x que no esta en X. Además se debe cumplir que:

$$\sum_{i=1}^{\infty} p(X_i) = 1$$

En este caso F(.) puede ser definida en términos de p(a) de la manera siguiente:

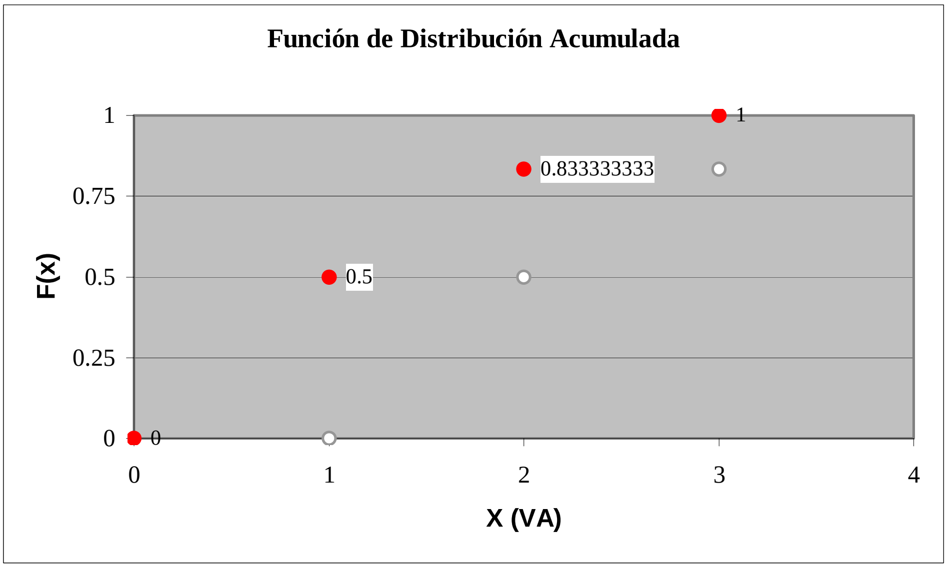
$$F(a) \ = \sum_{\forall x_i \le a} p(x_i)$$

**e1)** Supongamos que p(1)=1/2 ; p(2)=1/3 ; p(3)=1/6.




Luego la función acumulada de X es dado por:

$$F(a) \ = \begin{cases} 0 & si \ \ a < 1 \\ \frac{1}{2} & si \ \ 1 \le a < 2 \\ \frac{5}{6} & si \ \ 2 \le a < 3 \\ 1 & si \ \ a \ge 3 \end{cases}$$



**Para VA continuas** la relación entre la función de distribución acumulada F(.) y la función de densidad f(.) de una VA X continua es expresada a través de

$$F(a) \;=\; \int_{-\infty}^{a} f(t)\,dt$$

*y diferenciando a ambos lados tendremos:* $\dfrac{dF(a)}{d(a)} \;=\; f(a)$

(la derivada de la función de distribución acumulada es la función de densidad).

**Nota:**

$$\text{Debido a que}: P(\mathrm{a} < X \le b) = \int_a^b \mathrm{f}(x)\,\mathrm{d}\,x, \quad \text{tenemos}: \; P(b) = \int_b^b \mathrm{f}(x)\,\mathrm{d}\,x = 0 \quad y \text{ en consecuencia}:$$

$$P(\mathrm{a} < X < b) = P(\mathrm{a} \le X < b) = P(\mathrm{a} < X \le b) = P(\mathrm{a} \le X \le b)$$

Sin embargo, para mantener consistencia en la notación, utilizaremos $P(a < X \le b)$ como en el caso discreto.

## 3.2 Esperanza Matemática y Momentos

## Caso de VA Discretas:

Sea X una VA discreta finita que puede tomar valores $x_1, x_2, .. , x_n$, con las probabilidades $p(x_1), p(x_2), .. , p(x_n)$.

**Definición 1**
**Esperanza matemática o valor esperado (medio) de la VA X** discreta finita es la expresión:

$$E(X) = \sum_{i=1}^{n} x_i \, p(x_i)$$

**Definición 2**
**Esperanza matemática o valor medio de una función g(x)** de la VA X es la expresión:

$$E(g(x)) \;=\; \sum_{i=1}^{n} g(x_i)p(x_i)$$

**Ejemplos**

**e1)** El valor esperado de la VA X:= número que resulta al lanzar un dado, será:

$$E(X) \;=\; 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + \;...\; + 6 * \frac{1}{6} = \frac{1}{6} * (1 + 2 + ... + 6) = \frac{1}{6} * \frac{6(7)}{2} = 3.5$$

recordando: $\sum_{i=1}^{n} i = \dfrac{n(n+1)}{2}$

**e2)** Sea X: Suma de los números resultantes del lanzamiento de 2 dados, entonces,

$$E(X) = \frac{1}{36}(2 + 3 * 2 + 4 * 3 + 5 * 4 + ... + 10 * 3 + 11* 2 + 12) = \frac{1}{36} * (252) = 7$$

**e3)** Sea X: Número de divisores del número obtenido al sacar de una bolsa papeles numerados del 1 al 12.

| No. Extraído | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X | 1 | 2 | 2 | 3 | 2 | 4 | 2 | 4 | 3 | 4 | 2 | 6 |

| X | 1 | 2 | 3 | 4 | 6 |
|---|---|---|---|---|---|
| P | 1/12 | 5/12 | 1/6 | 1/4 | 1/12 |

Entonces,

$$E(X) = \frac{1}{12}(1 + 2.5 + 3.2 + 4.3 + 6.1) = 2.91$$

**Momentos de una VA.**

En muchas aplicaciones existe la necesidad de medir en alguna forma la dispersión de los valores de X alrededor de su valor medio E(X). Estos valores suelen ser esperanza matemática de ciertas funciones g(x), como por ejemplo:

**g(x) = | x - E(X) |**

Sin embargo, el manejo de valores absolutos es incómodo y por ello se prefieren potencias de X que dan lugar a los llamados momentos.

**Definición 1**

**Momento de X de grado (u orden) m** es el valor esperado de X elevado a la m, es decir:

$$\mu_m = E(X^m) = \sum_{i=1}^{n} x_i^m \, p(x_i)$$

En particular, $\mu_1 = \mu = E(X)$.

Los **momentos centrados** se definen a través de la ecuación:

$$\mu'_m = E((X - \mu)^m) = \sum_{i=1}^{n} (x_i - \mu)^m \, p(x_i)$$

En particular, como medida de dispersión, es importante el momento centrado de orden 2 definido a continuación.

**Definición 2**

Se conoce como **Varianza de una VA X** la siguiente expresión:

$$VAR(X) = \sigma^2(X) = \sigma^2 = E((X - \mu)^2) = \sum_{i=1}^{n} (x_i - \mu)^2 \, p(x_i)$$

El número no negativo σ, es decir, la raíz cuadrada de la Varianza se le conoce como la **Desviación Típica ó Standard** de la VA X.

Observación: Cuando la Varianza es grande ello indica que los valores de X están muy separados de su valor medio.

## Caso de VA Continuas:

Si X es una VA continua con la función de densidad *f(x)*, entonces:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

$$E(g(x)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

$$E(X^m) = \int_{-\infty}^{\infty} x^m f(x)dx$$

$$VAR(X) = E((X - E(X))^2)$$

**Lema 1.**
Con *a*, *b* constantes se cumple:

a) $E(aX + b) = aE(X) + b$

b) $E(X - \mu_1) = 0$

c) $\sigma^2(X) = E(X^2) - (E(X))^2$

d) $\sigma^2(aX + b) = a^2\sigma^2(X)$

(Fácilmente demostrables pero aplicando las respectivas definiciones)

*Demostración a)*

$$E(aX + b) = \sum_{i=1}^{n}(ax_i + b)p(x_i) = \sum_{i=1}^{n}(ax_i * p(x_i) + b * p(x_i)) = \sum_{i=1}^{n} ax_i * p(x_i) + \sum_{i=1}^{n} b * p(x_i) =$$

$$a\sum_{i=1}^{n} x_i p(x_i) + b\sum_{i=1}^{n} p(x_i) = aE(X) + b * (1) = aE(X) + b$$

*Demostración b)*

$$E(X - \mu_1) = \sum_{i=1}^{n}(x_i - \mu_1)p(x_i) = \sum_{i=1}^{n}(x_i * p(x_i) - \mu_1 * p(x_i)) = \sum_{i=1}^{n} x_i p(x_i) - \sum_{i=1}^{n}\mu_1 p(x_i) = E(x) - \mu_1\sum_{i=1}^{n} p(x_i) =$$

$$E(X) - \mu_1(1) = E(X) - E(X) = 0$$

*Demostración c)*

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

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

*Demostración d)*

$$\sigma^2(aX + b) =_{por\ (c)} E[(aX + b)^2] - [E(ax + b)]^2 = E(a^2X^2 + 2abX + b^2) - [aE(x) + b]^2 =$$

$$a^2E(X^2) + 2abE(X) + b^2 - [a^2(E(X))^2 + 2abE(X) + b^2] = a^2E(X^2) + 2abE(X) + b^2$$

$$- a^2(E(X))^2 - 2abE(X) - b^2 = a^2E(X)^2 - a^2(E(X))^2 = a^2[E(X)^2 - (E(X))^2] =_{por\ (c)} a^2\sigma^2(X)$$

## 3.3 Función generadora de Momentos

**Definición 1**
La **función generadora de momentos** se define como:

$$\phi(t) = E[e^{tX}] = \sum_{i=1}^{n} e^{tX}\, p(x_i) \ \ con X \ \ discreta$$
$$= \int_{-\infty}^{\infty} e^{tx}\, f(x)dx \ \ con\ X \ \ continua$$

Esta función se le conoce como generadora de momentos debido a que los momentos de X pueden ser obtenidos a través de sucesivas derivaciones de $\phi(t)$ y evaluando el resultado en $t=0$. Por ejemplo para obtener $E(X)$ se tiene:

$$\phi'(t) \ = \ \frac{d}{dt}E[\,e^{tX}\,] \ = \ E[\frac{d}{dt}e^{tX}\,] \ = \ E[\,Xe^{tX}\,]$$

Entonces, evaluando $\phi'(t)$ en $t=0$ tendremos: $\phi'(0) = E(X)$.
Igualmente:

$$\phi''(t) = \frac{d}{dt}\phi'(t) = \frac{d}{dt}E(\,Xe^{tX}\,) = E(\,X^2 e^{tX}\,) \quad con\, t=0,\ \phi''(0) = E(\,X^2\,)$$
$$........$$
$$\phi^n(t) = E(\,X^n e^{tX}\,),\ n \geq 1 \ \Rightarrow\ \phi^n(0) = E(\,X^n\,),\ n \geq 1$$

## 3.4 Desigualdad de Tchebycheff

Sea X una VA discreta y finita con valor medio $\mu$ y varianza $\sigma^2$, entonces

$$\forall k > 0,\ \ P(|x-\mu| \geq k) \ \leq \ \frac{\sigma^2}{k^2}.$$

Demostración.
Sea g(x) una función de X, tal que $g(x_i)>=0$, para todo $x_i$ en X, sea a > 0 constante, y sean $x_a$ aquellos valores tales que: $g(x_a) \geq a$. Entonces:

$$E(g(x)) \ = \ \sum_{i=1}^{n} g(\,x_i\,)p(\,x_i\,) \ \geq \ \sum_{x_a} g(\,x_a\,)p(\,x_a\,) \ \geq \ a\sum_{x_a} p(\,x_a\,)$$

Con el hecho siguiente,

$$\sum_{x_a} p(\,x_a\,) := \ \ Probabilidad\ de\ que \ \ \ g(x) \ \geq \ a$$

podemos concluir que,

$$E(g(x)) \geq a\, P\{g(x) \geq a\} \Rightarrow P\{g(x) \geq a\} \leq \frac{E(g(x))}{a}$$

En particular, con g(x):= x tendremos la **Desigualdad de Markov**

$$P\{X \geq a\} \leq \frac{E(X)}{a}$$

En el caso de Tchebychef obtendremos la desigualdad con $g(x) = (x - \mu)^2$ y $a = k^2$ tendremos,

$$P\{(x - \mu)^2 \geq k^2\} \leq \frac{E((x-\mu)^2)}{k^2}$$

por ello,

$$Prob\,\{|x - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

<u>Observación.</u>
La importancia de las desigualdades de Markov y Tchebychef reside en el hecho de que solo con el conocimiento de $\mu$ ( y de $\sigma$ en el caso de Tchebycheff) se puede acotar la probabilidad deseada. Sin embargo, esta cota no resulta buena es decir, informativa en la mayoría de los casos y por ello son usadas más frecuentemente en demostraciones de otras hipótesis.

**<u>Ejemplos</u>**

**e1)** En una fabrica de componentes, la producción semanal es aleatoria con un valor medio igual a 50 componentes. Que se puede decir acerca de la probabilidad de que la producción exceda las 75 componentes en una semana.

<u>Solución:</u>

Al aplicar la desigualdad de Markov se tiene:

$$P(X > 75) \leq \frac{E(X)}{75} = \frac{50}{75} = \frac{2}{3}$$

**e2)** Si la varianza de la producción semanal del problema anterior es igual a 25 que se puede decir acerca de la probabilidad de que la producción en una semana más de 40 y menos de 60 componentes.

<u>Solución:</u>

Aplicando la desigualdad de Tchebychef tendremos:

$$P\{40 < X < 60\} = P\{|X - 50| < 10\} = 1 - P\{|X - 50| \geq 10\}$$

$$P\,\{|x - 50| \geq 10\} \leq \frac{\sigma^2}{10^2} = \frac{1}{4}$$

Entonces,

$$P\{|x - 50| < 10\} \geq 1 - \frac{1}{4} = \frac{3}{4} \quad (75\%)$$

## 3.5 Transformación de Variables Aleatorias

**VA de tipo discreto:**

Sea *p(x)* la función de densidad probabílistica de una VA X. Si otra VA Y es obtenida de la VA X con una transformación Y = U(X), cuya inversa es X= W(Y). Entonces la función de densidad de Y viene dad por la siguiente expresión:

$$g(y) = \begin{cases} p(W(y)) & y \; en \; Y \\ 0 & en \; caso \; contrario \end{cases}$$

**Ejemplo.**

Sea X una VA con la siguiente función de densidad:

$$p(x) = \begin{cases} \mu^x \dfrac{e^{-\mu}}{x!} & x = 0,1,2,3,.. \\ 0 & \text{en caso contrario} \end{cases}$$

Sea la transformación y = 4x [y = U(x)], luego la VA Y = {0,4,8,12,..}.

Note que la transformación y = 4x establece una correspondencia uno a uno entre los espacios X e Y.
Nuestro objetivo consiste en hallar la f.d.p. de la VA Y, digamos g(y)=P[Y = y].

Obsérvese que la realización del evento Y = y ó 4X=y se da sii el evento X=(1/4)*y ocurre. En consecuencia:

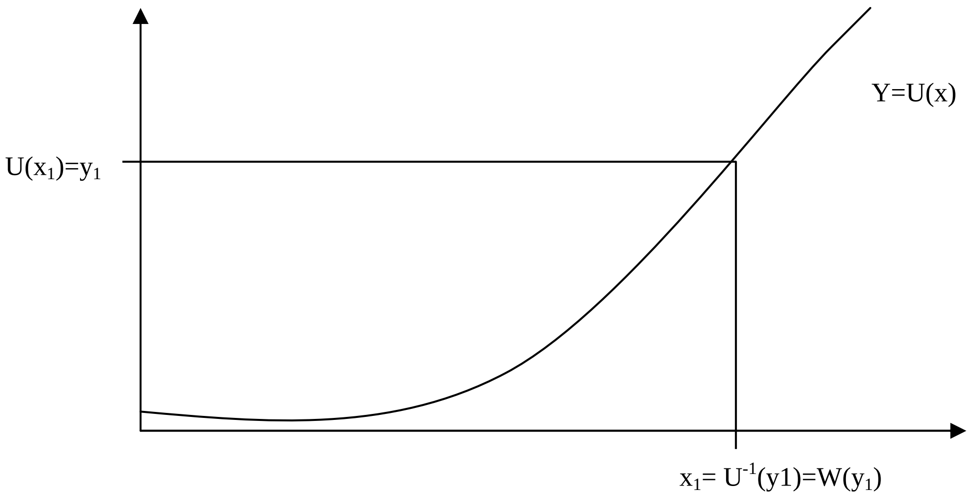
$$g(y) = P(Y = y) = P(4X = y) = P(X = \frac{y}{4}) = \frac{\mu^{\frac{y}{4}} e^{-\mu}}{(\frac{y}{4})!} \quad y = 0,4,8,\ldots$$

$$g(y) = \begin{cases} \mu^{\frac{y}{4}} \left( \dfrac{e^{-\mu}}{(\frac{y}{4})!} \right) & y = 0,4,8.. \\ 0 & en \; caso \; contrario \end{cases}$$

**VA de tipo continuo:**

Sea f(x) una f.d.p. de una VA X. Si otra VA Y es obtenida de X con la transformación Y = U(X) (diferenciable para todo x donde f(x)≠0) entonces la ecuación Y = U(X) puede resolverse de manera única para x con el fin de producir X= W(Y) (la inversa de Y = U(X)), y en consecuencia la función de densidad de Y viene dada por la siguiente expresión:

$$g(y) = \begin{cases} f(W(y)).|W'(y)| & y \; en \; Y \\ 0 & en \; caso \; contrario \end{cases}$$




Para mostrar la afirmación anterior con ayuda gráfica, observe que:

$\{x / U(x) \leq y_1\} = \{x / x \leq W(y_1)\}$ [ya que $\{x / U(x) \leq y_1\} = \{x / x \leq U^{-1}(y_1)\}$ ]

Ahora, debido a que $U(x)=Y$, $W(y_1)=x_1$ , tendremos $P(Y \leq y) = P(U(X) \leq y) = P(X \leq W(y)) = P(X \leq x)$,

donde $X=U^{-1}(Y)$ y en consecuencia:

$$F_Y(y) = F_X(x)$$

por lo tanto:

$$f_Y(y) = g(y) = dF_Y(y)/dy = dF_X(x)/dx = f(W(y)).|W'(y)|$$

**Ejemplo.**
Sea X una V.A. con función de densidad,

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Sea $Y = 8X^3$, luego la V.A. $Y = \{ y / 0< y <8 \}$

Sea $0< a < b < 8$, luego el evento $a < y < b$ ocurre sii el evento $a<8x^3<b$ ocurre, es decir: $(a/8)^{1/3} < x < (b/8)^{1/3}$ ;

$$P(a < Y < b) = P\left(\frac{1}{2}\sqrt[3]{a} < X < \frac{1}{2}\sqrt[3]{b}\right) = \int_{\frac{1}{2}\sqrt[3]{a}}^{\frac{1}{2}\sqrt[3]{b}} 2x\,dx$$

Ahora, si quisiéramos conocer f(y) o P(a<Y<b), tenemos que hacer las siguientes consideraciones:

Como:

$$y = U(x); \quad W(y) = \frac{1}{2} y^{\frac{1}{2}}$$

$$y = 8x^3 \quad \equiv> \quad x = \frac{1}{2} y^{\frac{1}{3}}$$

$$dx = \frac{1}{6} y^{-\frac{2}{3}} dy$$

Considerando los extremos de la integral:

$$Si \ x = \frac{1}{2}\sqrt[3]{b} \ => y = b, \quad si \ x = \frac{1}{2}\sqrt[3]{a} \ => y = a$$

$$P(a < Y < b) = \int_a^b 2(\frac{1}{2} y^{\frac{1}{3}}) \frac{1}{6} y^{-\frac{2}{3}} dy = \int_a^b \frac{y^{-\frac{1}{3}} dy}{6}$$

Es decir:

$$g(y) = \begin{cases} \dfrac{1}{6\sqrt[3]{y}} & 0 < y < 8 \\ 0 & en\ caso\ contrario \end{cases}$$

Observe que Y=U(x)=8x$^3$ y en consecuencia X=W(y)=1/2y$^{1/3}$, de donde se puede derivar que dx = W'(y)dy