3 - VARIABLES ALEATORIAS

3.1- Generalidades

En muchas situaciones experimentales se quiere asignar un número real a cada uno de los elementos del espacio muestral. Al describir el espacio muestral de un experimento un resultado individual no tiene que ser necesariamente un número, por ejemplo, al tirar una moneda y tomar como espacio muestral $S = \{c, s\}$, o al tirar un dado dos veces tomamos como espacio muestral a $S = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$, aquí S es un conjunto de pares ordenados.

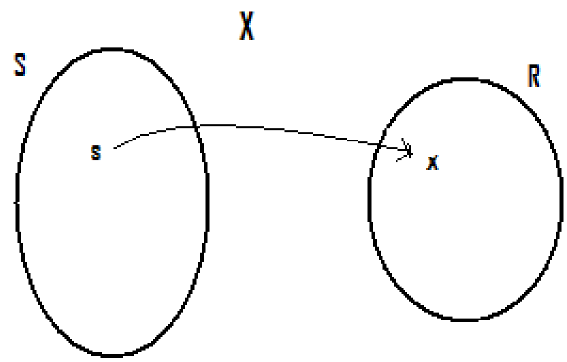
> Definición: Sea ε un experimento aleatorio y S un espacio muestral asociado a él. Una ***variable aleatoria*** es una función que asigna a cada elemento de S un número real.

Notación: se anota a una variable aleatoria con letras mayúsculas *X, Y, Z, W,…*

Entonces, si X es una variable aleatoria de S en R

$$X : S \to R \quad \text{tal que} \quad X(s) = x$$

Con diagramas de Venn



Desde ahora en lugar de escribir variable aleatoria, escribiremos v.a.

Ejemplos:

1- Se tira una moneda tres veces

Sea X la v.a. X: "número de caras obtenidas luego de los tres tiros"

Si tomamos como espacio muestral

$$S = \{(c,c,c); (c,c,s); (c,s,c); (s,c,c); (c,s,s); (s,s,c); (s,c,s); (s,s,s)\}$$

entonces

$$X((c,c,c)) = 3$$

$$X((c,c,s)) = X((s,c,c)) = X((c,s,c)) = 2$$

$$X((c,s,s)) = X((s,c,s)) = X((s,s,c)) = 1$$

$$X((s,s,s)) = 0$$

La imagen de esta función es el conjunto $\{0,1,2,3\}$

Dada una v.a. X a su imagen se la anota R_X y se la denomina ***rango o recorrido de X***

En el ejemplo anterior $R_X = \{0,1,2,3\}$

2- Se tira un dado tantas veces como sean necesarias hasta que sale el número 1 por primera vez.

Podemos simbolizar el espacio muestral de la siguiente manera $S = \{1,\ 01,\ 001,\ 0001,\ \ldots\}$, por ejemplo 001 simboliza el resultado que en los dos primeros tiros no salió el número 1 y en el tercer tiro salió el 1.
Sea Y la v.a.:
Y: "número de tiros necesarios hasta que sale el 1 por primera vez"
Entonces $R_Y = \{1,2,3,4,\ldots\}$, es decir el rango de Y es el conjunto de los números naturales.

3- En el interior de un círculo de radio r y centro el origen de coordenadas, se elige un punto al azar.
Tomamos como espacio muestral a $S = \{(x, y),\ \ x^2 + y^2 \le r^2\}$. Aquí S es infinito no numerable
Definimos la v.a. Z: "distancia del punto elegido al origen"
Entonces $R_Z = \{z;\ \ 0 \le z \le r\}$

Las variables aleatorias se clasifican según su rango.
Sea X es una v.a. con rango R_X. Si R_X es un conjunto ***finito o infinito numerable*** entonces se dice que X es una ***v.a. discreta***. Si R_X es un conjunto ***infinito no numerable*** entonces X es una ***v.a. continua***.

El rango R_X es considerado un nuevo ***espacio muestral***, y sus subconjuntos son ***eventos***.
Por ejemplo:
En el ejemplo 1, los eventos unitarios o elementales son $\{0\};\ \{1\};\ \{2\};\ \{3\}$, ***pero los anotamos*** $\{X = 0\};\ \{X = 1\};\ \{X = 2\};\ \{X = 3\}$
Otros eventos son, por ejemplo:
$\{X \le 1\}$, es decir, salió a lo sumo una cara.
Notar que podemos escribir $\{X \le 1\} = \{X = 0\} \cup \{X = 1\}$, o sea escribimos al evento como ***unión de eventos elementales***
$\{X > 0\}$, es decir, salieron una o más caras. Tenemos que $\{X > 0\} = R_X - \{X = 0\}$
En el ejemplo 2, $\{Y \ge 4\}$ sería el evento "al menos 4 tiros son necesarios para que salga por primera vez el numero 1"
$\{4 \le Y \le 6\}$ sería el evento "se necesitan entre 4 y 6 tiros para que salga el 1 por primera vez"
Notar que $\{4 \le Y \le 6\} = \{Y = 4\} \cup \{Y = 5\} \cup \{Y = 6\}$
En el ejemplo 3, $\left\{\frac{1}{3}r < Z < \frac{2}{3}r\right\}$ sería el evento "el punto elegido se encuentra a una distancia del centro mayor que $\frac{1}{3}r$, pero menor que $\frac{2}{3}r$

Volviendo al ejemplo 1, notar que $B = \{X = 0\}$ ocurre en R_X si y solo si el evento $A = \{(s,s,s)\}$ ocurre en S. Se dice que A y B son ***eventos equivalentes***.
De la misma forma los eventos
$A = \{(c,c,c)\}$ y $B = \{X = 3\}$ son equivalentes
$A = \{(c,c,s);(c,s,c);(s,c,c)\}$ y $B = \{X = 2\}$ son equivalentes
En general

siendo $A \subset S$ y $B \subset R_X$, A y B ***son equivalentes*** si $A = \{s \in S;\ \ X(s) \in B\}$

Si X es una v.a. de S en R, y R_X es el rango de X, para calcular la probabilidad de un evento B de R_X se busca el evento A en S equivalente a B y entonces $P(B) = P(A)$

Por ejemplo,

En el ejemplo 1, $P(B) = P(X = 0) = P(A) = P(\{(s,s,s)\}) = \dfrac{1}{8}$ si la moneda es normal.

$P(B) = P(X \le 1) = P(\{(s,s,s);(c,s,s);(s,c,s);(s,s,c)\}) = \dfrac{4}{8} = 0.5$ si la moneda es normal

También podríamos haber planteado

$$P(B) = P(X \le 1) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = 0.5$$

En el ejemplo 3, si $B = \left\{\dfrac{1}{3} r \le Z \le \dfrac{2}{3} r\right\}$, entonces B es equivalente a

$A = \left\{(x,y);\ \dfrac{1}{3} r \le \sqrt{x^2 + y^2} \le \dfrac{2}{3} r\right\}$, por lo tanto

$$P(B) = P(A) = \frac{\text{area de } A}{\text{area de } S} = \frac{\pi\left(\frac{2}{3} r\right)^2 - \pi\left(\frac{1}{3} r\right)^2}{\pi r^2} = \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 = \frac{3}{9} = \frac{1}{3}$$

<u>Observación</u>: en este ejemplo si $B = \left\{Z = \dfrac{2}{3} r\right\}$, entonces $P(B) = P(A) = \dfrac{\text{area de } A}{\text{area de } S} = \dfrac{0}{\pi r^2} = 0$

3.2 - <u>Variables aleatorias discretas</u>

Sea X una v.a. discreta. Anotamos su rango como $R_X = \{x_1, x_2, \dots, x_n\}$ si el rango es un conjunto finito de n elementos, y anotamos $R_X = \{x_1, x_2, \dots\dots\}$ si el rango es un conjunto infinito numerable.

A cada x_i se le asigna un número $p(x_i) = P(X = x_i)$. Estos números deben satisfacer las condiciones siguientes

a) $p(x_i) \ge 0$ para todo i

b) $\sum_i p(x_i) = 1$

La función $p(x)$ que antes se definió, se llama ***función de probabilidad o de frecuencia de la v.a. X***.
El conjunto de pares $(x_i, p(x_i))$ $i = 1,2,...$ es la ***distribución de probabilidad de X.***

Por ejemplo

1-Se tira una moneda normal tres veces, sea la v.a. X: "número de caras obtenidas"
Entonces $R_X = \{0,1,2,3\}$

Para hallar la distribución de probabilidad de X supongamos que la probabilidad de salir cara es ½ entonces

$$P(X = 0) = P(\{(s,s,s)\}) = \frac{1}{8}$$

$$P(X=1)=P(\{(c,s,s);(s,c,s);(s,s,c)\})=\frac{3}{8}$$

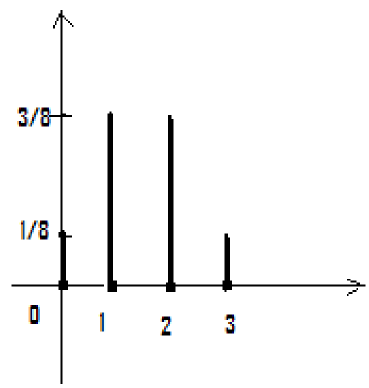
$$P(X=2)=P(\{(c,c,s);(s,c,c);(c,s,c)\})=\frac{3}{8}$$

$$P(X=3)=P(\{(c,c,c)\})=\frac{1}{8}$$

Se puede presentar la distribución de probabilidad de *X* en una tabla de la siguiente forma

x	0	1	2	3
p(x)	1/8	3/8	3/8	1/8

Un gráfico de la distribución de probabilidad de *X* sería



2-Se tira un dado normal. Sea X: "número que queda en la cara superior" Entonces $R_X=\{1,2,3,4,5,6\}$

La función de distribución de *X* es

x	1	2	3	4	5	6
p(x)	1/6	1/6	1/6	1/6	1/6	1/6

Observación:

Sea X una v.a. discreta con rango finito $R_X=\{x_1,x_2,\ldots,x_n\}$, donde cada x_i es un número entero y x_{i+1} es el consecutivo de x_i .

Si $P(X=x_i)=\frac{1}{n}$ para cada i entonces se dice que X tiene ***distribución uniforme discreta***.

Por ejemplo podría ser $R_X=\{1,2,3,\ldots,n-1,n\}$, en este caso X es uniforme discreta en el intervalo natural $[1,n]$.

La v.a. del ejemplo 2 anterior es uniforme discreta en el intervalo natural $[1,6]$

Función de distribución acumulada

Sea X una v.a. con rango R_X. Se define la ***función de distribución acumulada de X*** (abreviamos F.d.a de *X*) como

$$F(x)=P(X\leq x) \qquad -\infty<x<\infty \tag{12}$$

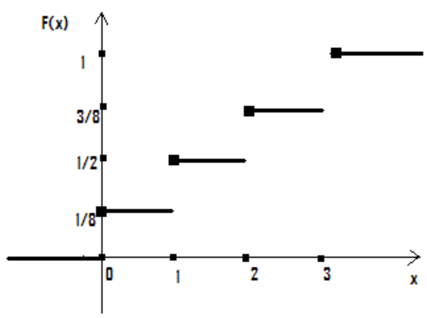
En el caso de ser X una v.a. discreta

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p(x_i) \qquad -\infty < x < \infty$$

Volviendo al ejemplo 1 anterior, la F.d.a. de X es

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{8} + \frac{3}{8} & \text{si } 1 \leq x < 2 \\ \frac{1}{8} + \frac{3}{8} + \frac{3}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x > 3 \end{cases} \quad \Rightarrow \quad F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{2} & \text{si } 1 \leq x < 2 \\ \frac{7}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x > 3 \end{cases}$$

La gráfica de la F.d.a. de X es



<u>Observación</u>: la F.d.a. de X es una función ***escalonada***, los puntos de "salto" coinciden con los puntos del rango de X, y la magnitud del salto en x_i es igual a $P(X = x_i)$

En general si X es una v.a. discreta cualquiera, su F.d.a. será una función escalonada.

Además si $x_1, x_2, ..., x_n$ son los valores del rango de X ordenados de menor a mayor entonces

$$P(X = x_1) = F(x_1)$$
$$P(X = x_i) = F(x_i) - F(x_{i-1}) \qquad i = 2, ..., n$$

Es decir, se puede obtener la función de distribución de X a partir de su F.d.a.

Para números cualesquiera a y b

1- Si $a \leq b$ entonces $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$

2- Si $a < b$ entonces $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$

3- Si $a < b$ entonces $F(a) \leq F(b)$ (es decir $F(x)$ es una función creciente)

Además se cumple que

$$1-\ \lim_{x\to\infty} F(x) = \lim_{x\to\infty} \sum_{x_i \le x} P(X = x_i) = \sum_{x_i} P(X = x_i) = 1$$

$$2-\ \lim_{x\to -\infty} F(x) = \lim_{x\to -\infty} \sum_{x_i \le x} P(X = x_i) = \sum_{x_i \le -\infty} P(X = x_i) = 0$$

3.3 – Esperanza de una variable aleatoria discreta

Sea X una v.a. discreta con rango R_X. La ***esperanza , valor medio o valor esperado de X***, lo anotamos $E(X)$, y se define como

$$E(X) = \sum_{x_i \in R_X} x_i P(X = x_i)$$

La sumatoria se hace sobre todos los posibles valores de X
Otra notación usual es μ_X o μ

Ejemplos:

1- Sea la v.a. X: "número que queda en la cara de arriba al tirar un dado normal" $R_X = \{1,2,3,4,5,6\}$

Entonces $E(X) = \sum_{x=1}^{6} xP(X = x) =$

$= 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) + 6P(X = 6) =$

$= 1\times\frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 6\times\frac{1}{6} = \frac{7}{2} = 3.5$

2- Se tira una moneda normal tres veces, sea la v.a. X: "número de caras obtenidas"

Entonces $R_X = \{0,1,2,3\}$

Calculamos la esperanza de X

$E(X) = \sum_{x=0}^{3} xP(X = x) = 0\times\frac{1}{8} + 1\times\frac{3}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8} = \frac{3}{2} = 1.5$

Observaciones:

1- La esperanza de una v.a. no tiene que coincidir necesariamente con algún valor del rango de la variable

2- En el ejemplo 1 donde el rango es finito y equiprobable, la esperanza de X coincide con el ***promedio de los valores del rango de X***

3- Se puede interpretar a la esperanza de una v.a. como un ***promedio "pesado" o "ponderado" de los valores del rango de la variable***, donde el "peso" de cada x_i es la probabilidad $P(X = x_i)$

4- Otra interpretación que se puede hacer de la esperanza es la siguiente: consideremos el ejemplo1, supongamos que tiramos el dado muchas veces, N veces, y entonces obtenemos una secuencia de N valores $x_1, x_2, \dots\dots, x_N$ donde cada x_i es un número natural del 1 al 6. Supongamos además que hacemos un promedio de esos N valores, y si llamamos n_i al número de veces que sale el número i tenemos que

$$\frac{x_1 + x_2 + \dots + x_N}{N} = \frac{n_1 1 + n_2 2 + \dots + n_6 6}{N} =$$

$$= \frac{n_1}{N}\times 1 + \frac{n_2}{N}\times 2 + \dots + \frac{n_6}{N}\times 6 \approx 1\times P(X=1) + 2\times P(X=2) + \dots + 6\times P(X=6) = E(X)$$

Es decir si promediamos los N valores medidos de X, ese promedio tiende a *E(X)* cuando $N \to \infty$, pues $\frac{n_i}{N} \approx P(X=i)$ cuando N es grande.

Esperanza de una función

A veces importa hallar la esperanza de una ***función de X*** y no de X misma. Veamos un ejemplo.
Un instructor de escritura técnica ha solicitado que cierto reporte sea entregado a la semana siguiente, agregando la restricción de que cualquier reporte que sobrepase las cuatro páginas será rechazado.
Sea X: "número de páginas del reporte de cierto estudiante seleccionado al azar"
Supongamos que X tenga la siguiente distribución de probabilidad

x	1	2	3	4
p(x)	0.01	0.19	0.35	0.45

Suponga que el instructor tarda $\sqrt{X}$ minutos calificando un trabajo que consiste en X páginas. Claramente $\sqrt{X}$ es ***otra variable aleatoria***. ¿Cuál será su esperanza?, es decir ¿a qué es igual $E\left(\sqrt{X}\right)$?
Para calcular la esperanza de una v.a. se necesita conocer su función de distribución de probabilidad, por lo tanto habría que hallar previamente la distribución de probabilidad de la v.a. $Y = \sqrt{X}$.
Está claro que si el rango de X es $R_X = \{1,2,3,4\}$ entonces el rango de Y será $R_Y = \{\sqrt{1}, \sqrt{2}, \sqrt{3}, \sqrt{4}\}$.
Además

$P(Y=\sqrt{1}) = P(X=1) = 0.01$

$P(Y=\sqrt{2}) = P(X=2) = 0.19$

$P(Y=\sqrt{3}) = P(X=3) = 0.35$

$P(Y=\sqrt{4}) = P(X=4) = 0.45$

Por lo tanto

$$E(Y) = \sqrt{1}\times P(Y=\sqrt{1}) + \sqrt{2}\times P(Y=\sqrt{2}) + \sqrt{3}\times P(Y=\sqrt{3}) + \sqrt{4}\times P(Y=\sqrt{4}) =$$

$$= \sqrt{1}\times P(X=1) + \sqrt{2}\times P(X=2) + \sqrt{3}\times P(X=3) + \sqrt{4}\times P(X=4) = 1.78491$$

O sea

$$E(Y) = \sum_x \sqrt{x}\,P(X=x)$$

Lo visto en este ejemplo se puede generalizar en el siguiente

> Teorema: Si X es una v.a. discreta con rango R_X y distribución de probabilidad *p(x)*, entonces la esperanza de cualquier función *h(X)* es igual a
>
> $$E(h(X)) = \sum_x h(x)\,p(x)$$

Ejemplo:
Un negocio de computadoras ha comprado tres computadoras de cierto tipo a \$500 cada una y las ven-

derá a \$1000 cada una. El fabricante ha aceptado volver a comprar en \$200 cualquier computadora que no se haya vendido en un tiempo especificado.
Sea X: "número de computadoras vendidas" , y supongamos que la distribución de probabilidad de X es

x	0	1	2	3
$p(x)$	0.1	0.2	0.3	0.4

Si consideramos la v.a. Y: "utilidad obtenida", entonces Y es una función de X, es decir $Y = h(X)$
Específicamente

$$Y = 1000X + 200(3 - X) - 1500 = 800X - 900$$

La utilidad esperada, es decir la $E(Y)$ será

$$E(Y) = \sum_{x=0}^{3} (800x - 900) P(X = x) =$$

$$= (800\times 0 - 900)P(X = 0)(800\times 1 - 900)P(X = 1) + (800\times 2 - 900)P(X = 2) + (800\times 3 - 900)P(X = 3) =$$

$$= (-900)\times 0.1 + (-100)\times 0.2 + 700\times 0.3 + 1500\times 0.4 = \$700$$

Notar que aplicando propiedades de la notación Σ se puede plantear

$$E(Y) = \sum_{x=0}^{3} (800x - 900) P(X = x) = 800\sum_{x=0}^{3} xP(X = x) - 900\sum_{x=0}^{3} P(X = x) = 800E(X) - 900$$

y calculando la esperanza de X , se llega al mismo resultado

<u>Propiedades de la esperanza</u>

En el ejemplo anterior tenemos que ***Y es una función lineal de X*** , es decir $Y = aX + b$ con a y b números reales.
En este caso vale entonces la siguiente propiedad

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

La demostración sigue los mismos pasos que en el ejemplo anterior

$$E(aX + b) = \sum_{x} (ax + b) P(X = x) = a \underbrace{\boxed{\sum_{x} xP(X = x)}}_{= E(X)} + b \underbrace{\boxed{\sum_{x} xP(X = x)}}_{= 1} = aE(X) + b$$

<u>Ejemplo</u>:
En el ejemplo anterior donde $Y = 800X - 900$
Directamente calculamos

$$E(Y) = 800E(X) - 900$$

Y

$$E(X) = 0\times 0.1 + 1\times 0.2 + 2\times 0.3 + 3\times 0.4 = 2$$

En consecuencia

$$E(Y) = 800E(X) - 900 = 800\times 2 - 900 = 700$$

Observaciones:
1- Para cualquier constante a, $E(aX) = aE(X)$
2- Para cualquier constante b, $E(X + b) = E(X) + b$

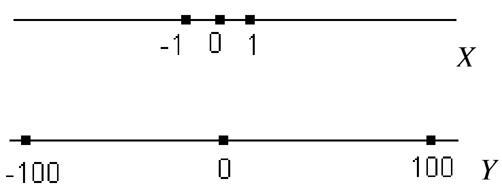
Varianza de una variable aleatoria

La esperanza de una v.a. mide dónde está centrada la distribución de probabilidad. Pero supongamos el siguiente ejemplo
Sean X e Y dos variables aleatorias con distribuciones dadas por

x	-1	1
p(x)	0.5	0.5

y	-100	100
p(y)	0.5	0.5

Es fácil verificar que $E(X) = E(Y) = 0$, pero los valores que toma la v.a. *Y* están más "alejados" de su esperanza que los valores de *X*.



Se busca una medida que refleje este hecho, se define entonces la varianza de una v.a.

Sea X una v.a. discreta con rango R_X , función de distribución de probabilidad *p(x)* y esperanza $E(X) = \mu$,
Entonces la ***varianza de X***, que anotamos $V(X)$, σ^2 o σ_X^2 es

$$V(X) = E\left[(X-\mu)^2\right] = \sum_{x \in R_X} (x-\mu)^2 p(x)$$

La ***desviación estándar de X*** es $\sigma_X = \sqrt{V(X)}$

Observaciones:
1- La varianza de una v.a. nunca es negativa
2- La cantidad $h(X) = (X-\mu)^2$ es el cuadrado de la desviación de *X* desde su media, y la varianza de *X* es la esperanza de la desviación al cuadrado. Si la mayor parte de la distribución de probabilidad está cerca de μ, entonces σ^2 será relativamente pequeña. Si hay valores de la variable alejados de μ que tengan alta probabilidad, entonces σ^2 será grande.
3- σ^2 está expresado en las unidades de medida de *X* al cuadrado, mientras que σ está expresada en las mismas unidades de medida que *X*.

Ejemplo:
En el caso de las variables aleatorias *X* e *Y* nombradas anteriormente,
$V(X) = (-1-0)^2 \times 0.5 + (1-0)^2 \times 0.5 = 1$ y $\sigma_X = 1$

$V(Y)=(-100-0)^2\times 0.5+(100-0)^2\times 0.5=100^2$ y $\sigma_Y=100$

Otra forma de escribir la varianza de una v.a., que facilita los cálculos es
$V(X)=\sum_{x\in R_X}(x-\mu)^2 p(x)=\sum_{x\in R_X}(x^2-2\mu x+\mu^2)p(x)=\sum_{x\in R_X}x^2 p(x)-2\mu\sum_{x\in R_X}xp(x)+\mu^2\sum_{x\in R_X}p(x)=$
$=E(X^2)-2\mu E(X)+\mu^2=E(X^2)-2\mu^2+\mu^2=E(X^2)-\mu^2$
Por lo tanto

$$V(X)=E(X^2)-\mu^2$$

Propiedades de la varianza

Las propiedades de la varianza de una v.a. son consecuencia de las propiedades de la esperanza de una v.a.
Si *X* es una v.a. discreta con rango R_X y distribución de probabilidad *p(x)*, entonces la varianza de cualquier función *h(X)* es igual a

$$V(h(X))=\sum_{x\in R_X}\left(h(x)-E(h(X))\right)^2 p(x)$$

Si *h(X)* es una función lineal, entonces

$$V(aX+b)=a^2V(X) \quad \text{y} \quad \sigma_{aX+b}=\sqrt{V(aX+b)}=|a|\sigma_X$$

Observaciones:
1- $V(aX)=a^2V(X)$
2- $V(X+b)=V(X)$

Ejemplo:
En un ejemplo anterior donde *X*: "número de computadoras vendidas" y *Y*: "utilidad obtenida", la *V(Y)* sería $V(Y)=800^2V(X)$
Necesitamos calcular *V(X)*
$V(X)=E(X^2)-\mu^2$

Sabemos ya que $\mu=E(X)=2$

Calculamos $E(X^2)=0^2\times 0.1+1^2\times 0.2+2^2\times 0.3+3^2\times 0.4=5$

En consecuencia

$V(Y)=800^2V(X)=800^2\left(E(X^2)-\mu^2\right)=800^2\left(5-2^2\right)=800$