

## II.1. Variable Aleatoria Discreta

March 31, 2019

### Variable Aleatoria Discreta

Consideremos que la variable aleatoria  $X$  sólo toma valores sobre un conjunto (finito o infinito) numerable de números. Los números Naturales,  $\mathbb{N}$ , los números Enteros,  $\mathbb{Z}$ , y los números Racionales,  $\mathbb{Q}$ , son ejemplos de conjuntos infinitos numerables. Por su parte, los seis números de las caras de un dado constituyen un ejemplo de conjunto numerable finito.

Consideremos que  $x_i$  son los valores que puede tomar la variable aleatoria, donde  $i \in I$  se refiere a un índice numerable que recorre  $I$  para poder indexar todos los valores. Por tanto  $\{X = x_i\}$  se refiere a un suceso elemental cualquiera.

Los sucesos correspondientes a cada valor de una variable discreta pueden tener probabilidades arbitrarias, que interpretaremos con el símil mecánico de *masa*. Podemos entonces definir una **función de masa de probabilidad (fmp)** que asigna una masa de probabilidad a cada suceso elemental representado por un valor posible de la variable aleatoria discreta

$$p_X(x_i) = P(\{X = x_i\}) \quad i \in I$$

Considerando las propiedades de la probabilidad, se cumple:

$$0 \leq p_X(x_i) \leq 1 \quad \forall i \in I$$

$$\sum_{i \in I} p_X(x_i) = 1$$

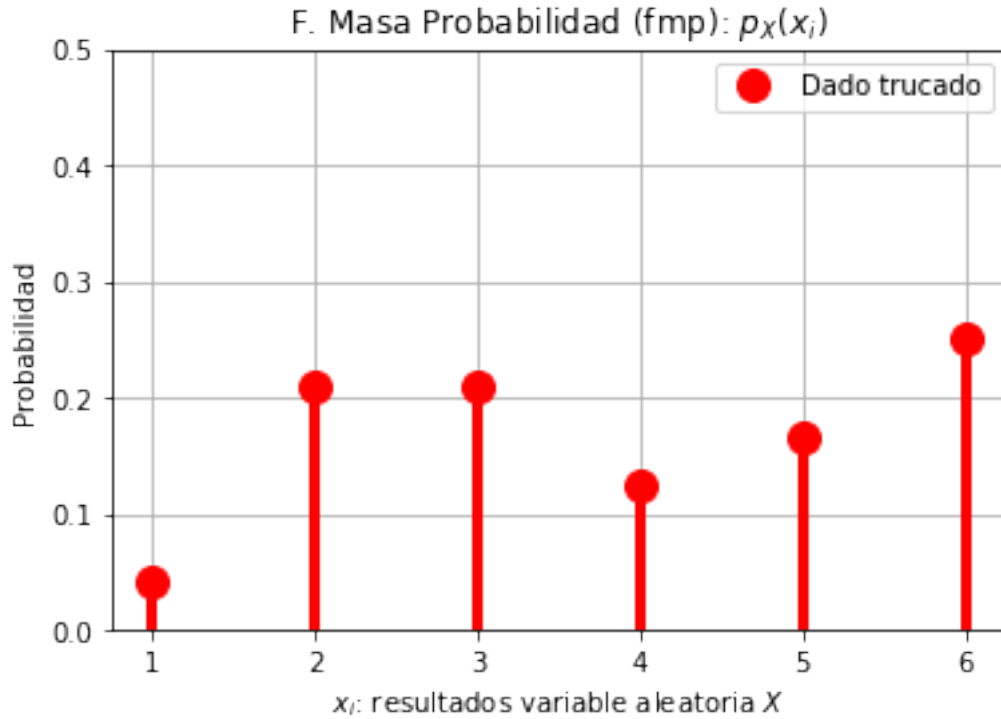
Por coherencia con la definición de variable aleatoria podemos considerar que  $p_X(x)$  está definida en todo el eje real, tomando **valor cero** fuera de los puntos  $x_i$  en los que hay masa de probabilidad.

Supongamos un dado trucado, cuya función de masa de probabilidad se representa en la siguiente tabla:

$X$	1	2	3	4	5	6	
$p_X(x_i)$	$\frac{1}{24}$	$\frac{5}{24}$	$\frac{5}{24}$	$\frac{3}{24}$	$\frac{4}{24}$	$\frac{6}{24}$	$\sum_i p_X(x_i) = 1$

Podemos fácilmente obtener las probabilidades de sucesos arbitrarios. Por ejemplo, la probabilidad de que salga 3 ó 5 es:

$$P(X = \{3, 5\}) = P(\{X = 3\} \cup \{X = 5\}) = p_X(3) + p_X(5) = \frac{9}{24}$$



Supongamos ahora, sin pérdida de generalidad, que el índice que numera los posibles valores de la variable aleatoria  $X$  es el conjunto de todos los números enteros  $\mathbb{Z}$ , y que la ordenación es tal que  $i < j \iff x_i < x_j$ .

Definimos la **función de probabilidad acumulada (FPA)**,  $F_X(x_i)$ , como sigue:

$$F_X(x_i) = \sum_{j=-\infty}^i p_X(x_j)$$

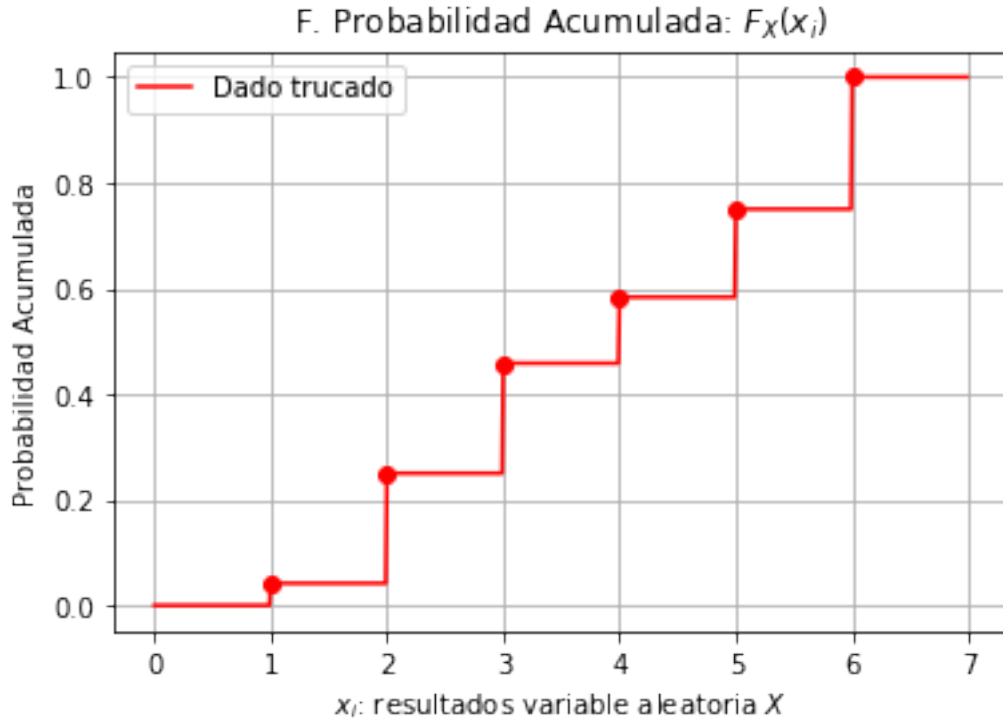
Adviértase que  $F_X(-\infty) = 0$ ,  $F_X(\infty) = 1$  y que  $F_X(x_i)$  es una función *no decreciente* en toda subsucesión de puntos  $x_i$ .

Por su parte, es fácil ver que  $p_X(x_i) = F_X(x_i) - F_X(x_{i-1})$

Calculemos la FPA para nuestro dado trucado, representándola en una tabla junto con la fmp:

$X$	0	1	2	3	4	5	6	7	
$p_X(x_i)$	0	$\frac{1}{24}$	$\frac{5}{24}$	$\frac{5}{24}$	$\frac{3}{24}$	$\frac{4}{24}$	$\frac{6}{24}$	0	$\sum_i p_X(x_i) = 1$
$F_X(x_i)$	0	$\frac{1}{24}$	$\frac{6}{24}$	$\frac{11}{24}$	$\frac{14}{24}$	$\frac{18}{24}$	$\frac{24}{24}$	$\frac{24}{24}$	$F_X(x_i \leq 0) = 0$ $F_X(x_i \geq 6) = 1$

Donde, para completar la representación hemos añadido los valores 0 y 7 que, lógicamente, corresponden a sucesos imposibles (como cualquier otro valor distinto de los primeros seis números naturales).



### Probabilidad condicionada

Considérese un suceso  $M$ , de probabilidad  $P(M)$ , del espacio muestral sobre el que se ha definido la variable aleatoria discreta  $X$ . Definimos la función de masa de probabilidad de  $X$  condicionada por  $M$  como sigue:

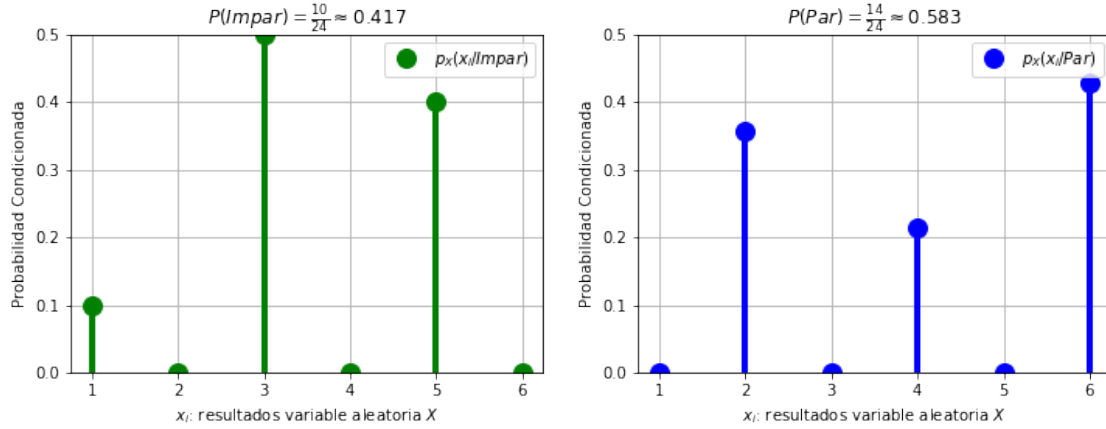
$$p_X(x_i/M) = \frac{P(\{X = x_i\} \cap M)}{P(M)}$$

Adviértase que  $\sum_i p_X(x_i/M) = 1$

Frecuentemente  $M$  podrá expresarse por medio de la variable aleatoria  $X$ , lo que simplifica el cálculo.

Consideremos los sucesos *Par* e *Impar* en el dado trucado que ya hemos visto y calculemos las funciones de masa de probabilidad condicionadas  $p_X(x_i/Par)$  y  $p_X(x_i/Impar)$

$X$	1	2	3	4	5	6	
$p_X(x_i)$	$\frac{1}{24}$	$\frac{5}{24}$	$\frac{5}{24}$	$\frac{3}{24}$	$\frac{4}{24}$	$\frac{6}{24}$	$\sum_i p_X(x_i) = 1$
$P(Impar)$	$\frac{1}{24}$	$+$	$\frac{5}{24}$	$+$	$\frac{4}{24}$		$P(Impar) = \frac{10}{24}$
$P(Par)$		$\frac{5}{24}$	$+$	$\frac{3}{24}$	$+$	$\frac{6}{24}$	$P(Par) = \frac{14}{24}$
$p_X(x_i/Impar)$	$\frac{1}{10}$	0	$\frac{5}{10}$	0	$\frac{4}{10}$	0	$\sum_i p_X(x_i/Impar) = 1$
$p_X(x_i/Par)$	0	$\frac{5}{14}$	0	$\frac{3}{14}$	0	$\frac{6}{14}$	$\sum_i p_X(x_i/Par) = 1$



Podemos también definir la probabilidad de un suceso  $M$  condicionada por un valor arbitrario  $x_i$  de la variable aleatoria  $X$ . Para ello no hay más que advertir que  $\{X = x_i\}$  es un suceso del espacio muestral:

$$P(M/X = x_i) = \frac{P(M \cap \{X = x_i\})}{P(\{X = x_i\})} = \frac{P(M \cap \{X = x_i\})}{p_X(x_i)}$$

Haremos uso frecuente del abuso de notación  $P(M/x_i) \equiv P(M/X = x_i)$ .

Adviértase que  $P(M/x_i) \neq p_X(x_i/M)$ , estando ambas expresiones relacionadas mediante el Teorema de Bayes, como se verá más adelante.

## Teorema de la Probabilidad Total

Consideremos una partición del espacio muestral

$$M_j \in \mathcal{P}(\Omega) \quad j = 1 \dots N \iff \begin{cases} \bigcup_{j=1}^N M_j = \Omega \\ M_j \cap M_k = \emptyset \quad j \neq k \end{cases}$$

y que conocemos las funciones de masa de probabilidad de una variable aleatoria  $X$  definida en el mismo condicionadas por todos los sucesos de dicha partición  $p_X(x_i/M_j)$ , así como las probabilidades de todos los sucesos condicionantes.

La función de masa de probabilidad total (sin condicionar) de la variable aleatoria  $X$  es

$$p_X(x_i) = \sum_{j=1}^N p_X(x_i/M_j)P(M_j)$$

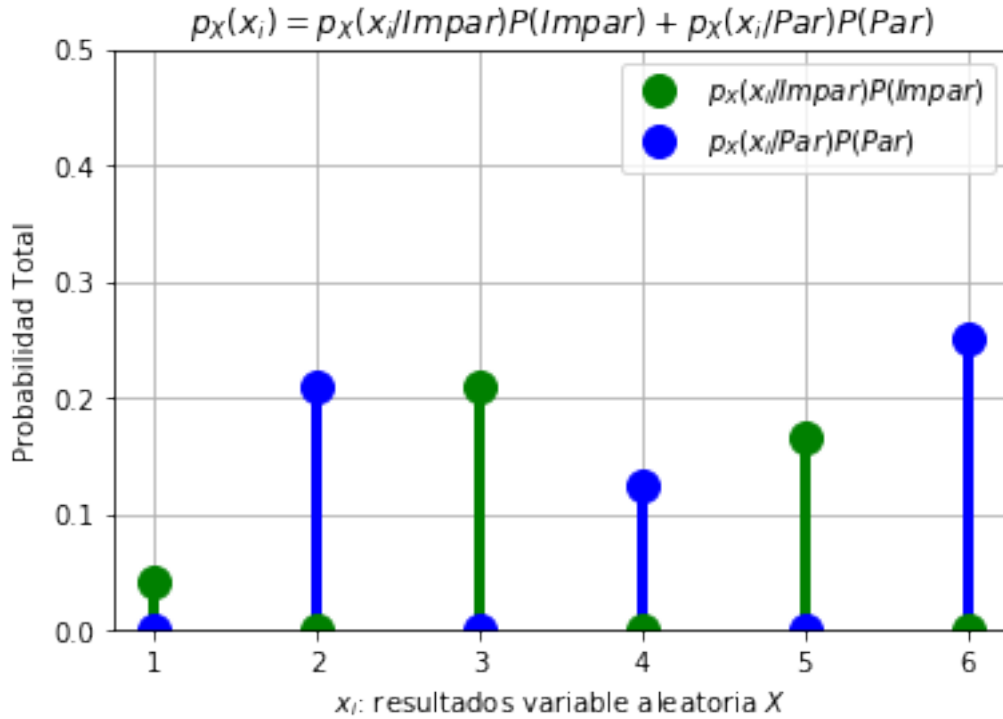
Adviértase que el teorema permite descomponer una función de masa de probabilidad en *mezcla* (mixture) de funciones de masa de probabilidad condicionadas, algo que resulta con frecuencia útil.

En el caso de tener un único suceso  $M$ , consideramos que el espacio muestral se particiona con él y con su suceso complementario  $\bar{M}$ , y que  $P(\bar{M}) = 1 - P(M)$ , resultando:

$$p_X(x_i) = p_X(x_i/M)P(M) + p_X(x_i/\bar{M})(1 - P(M))$$

En el ejemplo del dado truco podemos advertir fácilmente que

$$p_X(x_i) = p_X(x_i/Impar)P(Impar) + p_X(x_i/Par)P(Par)$$



La probabilidad total también puede aplicarse de forma alternativa, si lo que se conocen son las probabilidades un suceso arbitrario  $M$  condicionadas por cada uno de los valores posibles de la variable aleatoria  $X$ , esto es, si se conocen  $P(M/x_i) \equiv P(M/X = x_i)$ .

En tal caso, la probabilidad total del suceso  $M$  (sin condicionar) es

$$P(M) = \sum_{i=-\infty}^{\infty} P(M/x_i)p_X(x_i)$$

### Teorema de Bayes

Partiendo de la definición de función de masa de probabilidad condicionada es inmediato obtener el **Teorema de Bayes** (recordar que  $P(M/x_i) \equiv P(M/X = x_i)$ ):

$$p_X(x_i/M) = \frac{P(M/x_i)p_X(x_i)}{P(M)} \iff P(M/x_i) = \frac{p_X(x_i/M)P(M)}{p(x_i)}$$

Con frecuencia, los denominadores de las expresiones anteriores se obtienen mediante el *Teorema de la Probabilidad Total*, pues tan sólo se conocen las probabilidades condicionadas:

$$p_X(x_i/M) = \frac{P(M/x_i)p_X(x_i)}{\sum_{i=-\infty}^{\infty} P(M/x_i)p_X(x_i)}$$

$$P(M/x_i) = \frac{p_X(x_i/M)P(M)}{p_X(x_i/M)P(M) + p_X(x_i/\bar{M})(1 - P(M))}$$

Apliquemos el Teorema de Bayes al ejemplo del dado truco para obtener, trivialmente, las probabilidades de los sucesos *Par* e *Impar* condicionadas por los valores de la variable aleatoria:

X	1	2	3	4	5	6	
$p_X(x_i)$	$\frac{1}{24}$	$\frac{5}{24}$	$\frac{5}{24}$	$\frac{3}{24}$	$\frac{4}{24}$	$\frac{6}{24}$	$\sum_i p_X(x_i) = 1$
$P(\text{Impar})$	$\frac{1}{24}$	$+$	$\frac{5}{24}$	$+$	$\frac{4}{24}$		$P(\text{Impar}) = \frac{10}{24}$
$P(\text{Par})$		$\frac{5}{24}$	$+$	$\frac{3}{24}$	$+$	$\frac{6}{24}$	$P(\text{Par}) = \frac{14}{24}$
$p_X(x_i/\text{Impar})$	$\frac{1}{10}$	0	$\frac{5}{10}$	0	$\frac{4}{10}$	0	$\sum_i p_X(x_i/\text{Impar}) = 1$
$p_X(x_i/\text{Par})$	0	$\frac{5}{14}$	0	$\frac{3}{14}$	0	$\frac{6}{14}$	$\sum_i p_X(x_i/\text{Par}) = 1$
$P(\text{Impar}/x_i)$	1	0	1	0	1	0	$\sum_i p_X(\text{Impar}/x_i) \neq 1$
$P(\text{Par}/x_i)$	0	1	0	1	0	1	$\sum_i p_X(\text{Par}/x_i) \neq 1$

Podemos extender fácilmente el **Teorema de la Probabilidad Total** y el **Teorema de Bayes** considerando que los sucesos condicionantes  $M_j$  pueden modelarse con una variable aleatoria  $M$  cuya función de masa de probabilidad  $p_M(m_i)$  conocemos. Las funciones de masa de probabilidad condicionadas son  $p_{X/M}(x_i/m_j)$  y  $p_{M/X}(m_i/x_j)$ . En tal caso:

$$p_X(x_i) = \sum_j p_{X/M}(x_i/m_j)p_M(m_j)$$

$$p_{M/X}(m_i/x_j) = \frac{p_{X/M}(x_j/m_i)p_M(m_i)}{\sum_i p_{X/M}(x_j/m_i)p_M(m_i)}$$

## Función de variable aleatoria

Consideremos una variable aleatoria discreta  $X$ . Podemos aplicar sobre la misma una función  $g$  que devolverá valores aleatorios, al serlo también su variable independiente. Lo representamos como sigue:

$$Y = g(X)$$

Cada valor  $x_i$  de  $X$  se transformará por la función en un nuevo valor  $y_i$  de  $Y$ . Adviértase que valores distintos  $x_i \neq x_j$  pueden dar lugar a valores iguales  $y_i = y_j$  con tal que  $g(x_i) = g(x_j)$ . Por tanto, definimos la función de masa de probabilidad de  $Y$  como sigue:

$$p_Y(y_i) = \sum_j p_X(g^{-1}(y_i)) = \sum_{\substack{j \\ g(x_j)=y_i}} p_X(x_j)$$

Donde  $j$  indexa las raíces de la ecuación  $y_i = g(x_j)$ .

Consideremos, por ejemplo, la variable aleatoria  $X$  con la siguiente fmp:

X	-2	-1	0	1	
$p_X(x_i)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\sum_i p_X(x_i) = 1$

Veamos cuál es la fmp de la variable aleatoria  $Y = X^2$ :

$Y$	0	1	4
$p_Y(y_i)$	$\frac{1}{2}$	$\frac{1}{4} + \frac{1}{8}$	$\frac{1}{8}$
$\sum_i p_Y(y_i) = 1$			

Si se desean incluir otros valores de  $Y$  en la representación, por ejemplo el 2 ó el 3, tendrán masa de probabilidad nulas.

## Esperanza Matemática y Valor Medio. Momentos

Se define la esperanza matemática de una función  $g(X)$  de una variable aleatoria como sigue:

$$E(g(X)) = \sum_i g(x_i) p_X(x_i)$$

Es fácil advertir que el operador  $E()$  es lineal, esto es que

$$E(a_1 g_1(X) + a_2 g_2(X)) = a_1 E(g_1(X)) + a_2 E(g_2(X))$$

donde  $a_1$  y  $a_2$  son coeficiente constantes arbitrarios y  $g_1$  y  $g_2$  son dos funciones. También es evidente que, si  $k$  es una constante arbitraria

$$E(k) = k$$

Si  $g$  es la función identidad obtenemos el **valor medio** de la variable aleatoria  $X$ :

$$\eta_X = E(X) = \sum_i x_i p_X(x_i)$$

Es conveniente **reflexionar sobre la relación con la media muestral**, advirtiéndole que ambas son medidas de tendencia central.

Adviértase que, como una función de variable aleatoria define una nueva variable aleatoria,  $Y = g(X)$ , con su propia función de masa de probabilidad, tenemos dos maneras de calcular la esperanza de una función de variable aleatoria:

$$\eta_Y = E(Y) = \sum_i y_i p_Y(y_i) = \sum_i g(x_i) p_X(x_i)$$

Consideremos que la variable aleatoria  $X$  tiene valor medio  $\eta_X$ . Calculemos cuál será el valor medio de la variable aleatoria  $Y = aX + b$ , donde  $a$  y  $b$  son coeficientes constantes.

Aplicando linealidad:

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

Son interesantes los casos particulares:

- Sumar constante,  $a = 1$ :  $E(X + b) = \eta_X + b$
- Multiplicar por constante  $a$ ,  $b = 0$ :  $E(aX) = a\eta_X$

## Momentos de Segundo Orden: Valor Cuadrático Medio y Varianza

El **valor cuadrático medio**, o momento no centrado de segundo orden, de la variable aleatoria  $X$  se define:

$$E(X^2) = \sum_i x_i^2 p_X(x_i) \geq 0$$

El valor cuadrático medio es una medida de dispersión respecto del origen, a la que no afecta el signo positivo o negativo de las excursiones (al estar elevada al cuadrado) sino solo la amplitud de las mismas. Por ello siempre es una magnitud positiva, salvo en el caso trivial de que toda la masa de probabilidad se localice en el origen (el valor es determinista e igual a cero).

Es conveniente reflexionar sobre su relación con el estadístico muestral correspondiente.

Consideremos que la variable aleatoria  $X$  tiene valor medio  $\eta_X$  y valor cuadrático medio  $E(X^2)$ . Calculemos cuál será el valor cuadrático medio de la variable aleatoria  $Y = aX + b$ , donde  $a$  y  $b$  son coeficientes constantes.

Desarrollando el cuadrado y aplicando linealidad:

$$E((aX + b)^2) = a^2 E(X^2) + (2ab\eta_X + b^2)$$

Es interesante el caso de multiplicar por una constante  $a, b = 0$ :

$$E((aX)^2) = a^2 E(X^2)$$

La **varianza** o momento centrado de segundo orden, se define:

$$\sigma_X^2 \equiv \text{Var}(X) = E((X - \eta_X)^2) \geq 0$$

La varianza es una medida de dispersión de los valores de la variable aleatoria en relación a su valor medio. De nuevo, al estar elevada al cuadrado, no importa el signo de las excursiones, sino sólo su amplitud y, por ello, siempre es una magnitud positiva, o trivialmente nula si toda la masa de probabilidad está localizada en el mismo punto (valor determinista).

La raíz cuadrada de la varianza se denomina **desviación típica**,  $\sigma_X$ .

Desarrollando la definición y aplicando la propiedad de linealidad, se obtiene:

$$\sigma_X^2 = E(X^2) - \eta_X^2 \implies E(X^2) \geq \sigma_X^2$$

Si la variable aleatoria tiene **media nula**:  $E(X^2) = \sigma_X^2$

Consideremos ahora que la variable aleatoria  $X$  tiene valor medio  $\eta_X$ , valor cuadrático medio  $E(X^2)$  y varianza  $\sigma_X^2 \equiv \text{Var}(X)$ . Calculemos cuál será la varianza de la variable aleatoria  $Y = aX + b$ , donde  $a$  y  $b$  son coeficientes constantes.

Aplicando la relación entre valor cuadrático medio, varianza y cuadrado de la media:

$$\text{Var}(aX + b) = E((aX + b)^2) - E^2(aX + b)$$

Utilizando las expresiones que ya tenemos:

$$\begin{aligned} E((aX + b)^2) &= a^2 E(X^2) + 2ab\eta_X + b^2 \\ E^2(aX + b) &= (a\eta_X + b)^2 = a^2\eta_X^2 + 2ab\eta_X + b^2 \end{aligned}$$

Resultando la varianza:



$$\begin{aligned} \text{Var}(aX + b) &= a^2 E(X^2) - a^2 \eta_X^2 = a^2 (E(X^2) - \eta_X^2) \\ &= a^2 \text{Var}(X) \end{aligned}$$

Dado trucado:

Media: 3.916666666666667

Varianza: 2.576388888888884

Valor Cuadrático Medio: 17.916666666666664

### Tipificación de una Variable Aleatoria

Consideremos que una variable aleatoria  $X$  tiene valor medio  $\eta_X$  y varianza  $\sigma_X$ . **Tipificar** la variable quiere decir buscar otra equivalente, pero con **media nula** y **varianza unidad**.

La variable tipificada  $\hat{X}$  la obtenemos aplicando la siguiente función:

$$\hat{X} = \frac{X - \eta_X}{\sigma_X}$$

De lo visto anteriormente, es inmediato advertir que:

$$\begin{aligned} \eta_{\hat{X}} &= E(\hat{X}) = \frac{1}{\sigma_X} \eta_X - \frac{\eta_X}{\sigma_X} &= 0 \\ \sigma_{\hat{X}}^2 &= \text{Var}(\hat{X}) = \frac{1}{\sigma_X^2} \sigma_X^2 &= 1 \end{aligned}$$

### Momentos de orden superior

El momento no centrado de orden  $n$  se define como:

$$m_n = E(X^n) = \sum_i x_i^n p_X(x_i)$$

El momento centrado de orden  $n$ :

$$\mu_n = E((X - \eta_X)^n) = \sum_i (x_i - \eta_X)^n p_X(x_i)$$

Y el momento tipificado de orden  $n$ :

$$\gamma_n = E\left(\left(\frac{X - \eta_X}{\sigma_X}\right)^n\right) = \sum_i \left(\frac{x_i - \eta_X}{\sigma_X}\right)^n p_X(x_i)$$

El momento tipificado de orden 3 corresponde al coeficiente de asimetría, y el de orden 4 al coeficiente de curtosis. Es conveniente reflexionar sobre su relación con sus equivalentes muestrales.

## Esperanza Matemática Condicionada

Podemos definir la esperanza matemática de una función  $g$  de la variable aleatoria  $X$  condicionada por el suceso  $M$  como sigue:

$$E(g(X)/M) = \sum_i g(x_i) p_X(x_i/M)$$

Podemos extender con la esperanza condicionada las propiedades y definiciones vistas anteriormente. Por ejemplo:

$$\begin{aligned}\eta_{X/M} &= E(X/M) = \sum_i x_i p_X(x_i/M) \\ E(X^2/M) &= \sum_i x_i^2 p_X(x_i/M) \\ \sigma_{X/M}^2 &= \text{Var}(X/M) = \sum_i (x_i - \eta_{X/M})^2 p_X(x_i/M) = E(X^2/M) - \eta_{X/M}^2\end{aligned}$$

Dado trucado - Estadísticos condicionados por ser Impar:

Media condicionada por Impar: 3.6000000000000005

Varianza condicionada por Impar: 1.6399999999999997

Valor Cuadrático Medio condicionado por Impar: 14.600000000000001

Dado trucado - Estadísticos condicionados por ser Par:

Media condicionada por Par: 4.142857142857142

Varianza condicionada por Par: 3.122448979591841

Valor Cuadrático Medio condicionado por Par: 20.285714285714285