Variables aleatorias discretas

Al realizar un experimento generalmente estamos interesados en alguna función del resultado más que en el resultado en sí mismo. Así, por ejemplo, al arrojar un dado dos veces podríamos estar interesados sólo en la suma de los puntos obtenidos y no en el par de valores que dio origen a ese valor de la suma. Esa cantidad de interés, o más formalmente esa función a valores reales definida sobre el espacio muestral se denomina *variable aleatoria. Variable* porque toma distintos valores y *aleatoria* porque el valor observado no puede ser predicho antes de la realización del experimento, aunque sí se sabe cuáles son sus posibles valores.

Dado que el valor de una variable aleatoria (en adelante lo abreviaremos v.a.) es determinado por el resultado de un experimento, podremos asignar probabilidades a los posibles valores o conjuntos de valores de la variable.

Ejemplo: Se arroja dos veces un dado equilibrado. Un espacio muestral asociado es:

$$S = \{(x_1, x_2) \,/\, x_i \in \{1,2,3,4,5,6\}\}$$

Posibles v.a. asociadas con este experimento son:

X: "número de caras pares"
Y: "máximo puntaje"
Z: "suma de puntos"

Definición: Sea S un espacio muestral asociado con un experimento aleatorio. Una variable aleatoria *X* es una función que asocia a cada elemento $w \in S$ un número real *X(w)=x*, es decir

$$X : S \rightarrow \Re$$

Como se observa, en general representaremos a las v.a. con letras mayúsculas: *X, Y, Z*, etc. y sus valores con letras minúsculas, es decir *X(w)=x* significa que *x* es el número real asociado al resultado $w \in S$ a través de *X*.

Ejemplos: 1) Volviendo al ejemplo anterior,

X((2,5)) = 1	*Y((2,5)) = 5*	*Z((2,5)) = 7*
X((1,3)) = 0	*Y((1,3)) = 3*	*Z((1,3)) = 4*
X((2,2)) = 2	*Y((2,2)) = 2*	*Z((2,2)) = 4*

2) Se arroja una moneda equilibrada 3 veces,

$$X = \begin{cases} 1 & \text{si el número de caras es impar} \\ 0 & \text{en caso contrario} \end{cases}$$

3) Se arroja una moneda equilibrada hasta que se obtiene la primera cara,

X: “número de tiros necesarios”

4) A partir del instante en que se intenta la conexión a un servidor, se registra el tiempo que demora en concretarse la misma,

X: “tiempo requerido para la conexión”.

En los ejemplos 1), 2) y 3) las v.a. toman un número finito o infinito numerable de valores, mientras que en el ejemplo 4) la v.a. *X* toma valores en un conjunto infinito no numerable, el intervalo $(0, \infty)$ o un intervalo (0, M) si existe un tiempo máximo (“time out”).

Notación: Indicaremos con R_X el rango de la v.a. *X*, es decir el conjunto de valores posibles de la v.a. *X*.

Ejemplos: En los ejemplos anteriores,

1) $R_X = \{0,1,2\}$
   $R_Y = \{1,2,3,4,5,6\}$
   $R_Z = \{2,3,4,5,6,7,8,9,10,11,12\}$

2) $R_X = \{0,1\}$

3) $R_X = \{1,2,3,...\} = N$

4) $R_X = (0,\infty)$ ó (0,M) si existe un “time out”

Definición: Una v.a. es *discreta* si toma un número finito o infinito numerable de valores.

Ejemplo: En el caso del ejemplo 1), ¿cómo calcularíamos la probabilidad de que la v.a. *Z* tome el valor 7, suponiendo que los lanzamientos son independientes?

$$P(Z=7) = P\left(\left\{(x_1,x_2) \in S \,/\, Z\left((x_1,x_2)\right) = 7\right\}\right) = P\left(\left\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\right\}\right) = \frac{6}{36} = \frac{1}{6}.$$

Definición: La **función de probabilidad puntual o de masa** de la v.a. discreta *X*, se define para todo *x* como

$$p_X(x) = P(X = x) = P\left(\left\{w \in S \,/\, X(w) = x\right\}\right)$$

Se cumplen las siguientes propiedades:

$$p_X(x) \geq 0 \qquad \forall\, x$$

$$\sum_{x \in R_X} p_X(x) = 1$$

La función de probabilidad puntual de una v.a. *X* nos dice cómo se distribuye la probabilidad total entre los distintos valores de *X*, y se determina a partir de la probabilidad de los sucesos asociados a cada valor de *X*.

Ejemplos: 1) Hallemos la función de probabilidad puntual de la v.a. *X* : "número de caras pares al arrojar dos veces un dado equilibrado". Recordemos que $R_X = \{0,1,2\}$.

$$p_X(0) = P(X = 0) = P\left(\{(x_1, x_2) \in S \,/\, x_1, x_2 \in \{1,3,5\}\}\right) = \frac{9}{36} = \frac{1}{4}$$

$$p_X(1) = P(X = 1) =$$

$$= P\left(\{(x_1, x_2) \in S \,/\, x_1 \in \{1,3,5\}, x_2 \in \{2,4,6\}\} \cup \{(x_1, x_2) \in S \,/\, x_1 \in \{2,4,6\}, x_2 \in \{1,3,5\}\}\right) = \frac{18}{36} = \frac{1}{2}$$

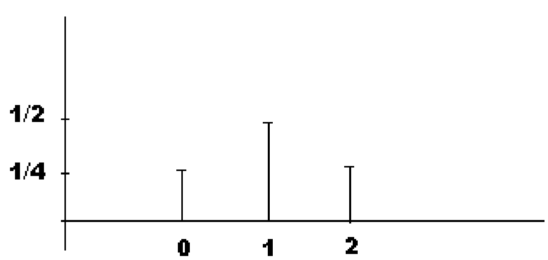
$$p_X(2) = P(X = 2) = P\{(x_1, x_2) \in S \,/\, x_1, x_2 \in \{2,4,6\}\} = \frac{9}{36} = \frac{1}{4}$$

Podemos resumir esta información en una tabla de la forma:

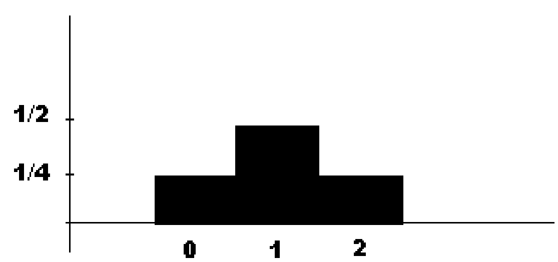
x	0	1	2
$p_X(x)$	1/4	1/2	1/4

o mediante un gráfico en el cual, para cada valor de *x* se construye una barra o un rectángulo centrado en x, cuya altura es proporcional a $p_X(x)$

Diagrama de Barras



Histograma



<u>Definición:</u> La **función de distribución acumulada** de una v.a. discreta *X* con función de probabilidad puntual $p_X(x)$ se define para todo $x \in \Re$, como

$$F_X(x) = P(X \le x) = \sum_{y \le x,\, y \in R_X} p_X(y)$$

Es decir que $F_X(x)$ es la probabilidad de que la v.a. *X* tome valores menores o iguales que *x*.

<u>Ejemplo</u>: Volviendo al ejemplo 1), hallemos la función de distribución acumulada de la v.a. *X,* cuya función de probabilidad puntual es

x	0	1	2
$p_X(x)$	1/4	1/2	1/4

Si $x < 0$ $\quad F_X(x) = P(X \le x) = 0$

$x = 0$ $\quad F_X(0) = P(X \le 0) = p_X(0) = 1/4$

$0 < x < 1$ $\quad F_X(x) = P(X \le x) = p_X(0) = 1/4$

$x = 1$ $\quad F_X(1) = P(X \le 1) = p_X(0) + p_X(1) = 1/4 + 1/2 = 3/4$

$1 < x < 2$ $\quad F_X(x) = P(X \le x) = p_X(0) + p_X(1) = 3/4$

$x = 2$ $\quad F_X(2) = P(X \le 2) = p_X(0) + p_X(1) + p_X(2) = 1/4 + 1/2 + 1/4 = 1$

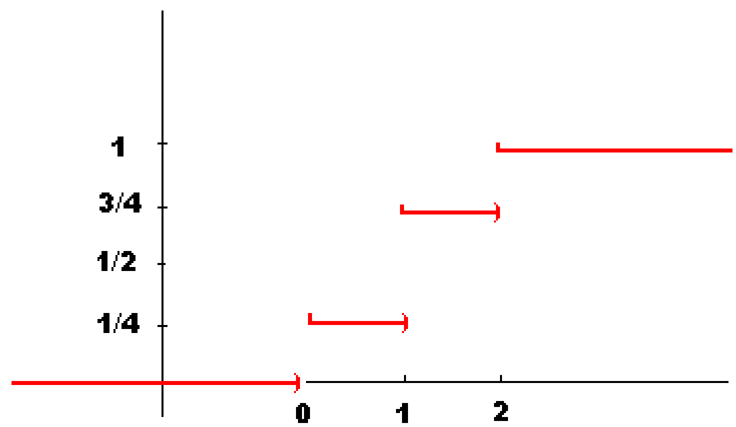
$x > 2$ $\quad F_X(x) = P(X \le 2) = p_X(0) + p_X(1) + p_X(2) = 1/4 + 1/2 + 1/4 = 1$

Resumiendo:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/4 & \text{si } 0 \le x < 1 \\ 3/4 & \text{si } 1 \le x < 2 \\ 1 & \text{si } x \ge 2 \end{cases}$$

¿Cómo es $F_X(x)$?

Observamos que se trata de una función escalera, no decreciente que toma valores entre 0 y 1.



Propiedades de la función de distribución acumulada:

i) $\forall x \in \Re,\ F_X(x) \in [0,1]$.

ii) $F_X(x)$ es monótona no decreciente, es decir que si $x_1 < x_2 \Rightarrow F_X(x_1) \le F_X(x_2)$.

iii) $F_X(x)$ es continua a derecha, es decir $\lim_{h \to o^+} F_X(x+h) = F_X(x)$.

iv) $\lim_{x \to \infty} F_X(x) = 1$ y $\lim_{x \to -\infty} F_X(x) = 0$

v) En cada punto *x,* el valor del salto es la probabilidad puntual, es decir

$$p_X(x) = F_X(x) - F_X(x^-)$$

donde $x^- = \lim_{h \to 0^+}(x-h)$ (límite por la izquierda). En particular si *X* toma valores $x_1 < x_2 < \ldots$, entonces $p_X(x_i) = F_X(x_i) - F_X(x_{i-1})$ para todo $i \ge 2$ y $p_X(x_1) = F_X(x_1)$.

Dem: Daremos sólo demostraciones heurísticas de estas propiedades. Demostraciones rigurosas pueden encontrarse, por ejemplo, en S. Ross (1988) o B. James (1981).

i) Obvio, ya que $F_X(x) = P(X \le x) = P(\{w \in S \,/\, X(s) \le x\})$ y toda probabilidad toma valores entre 0 y 1.

ii) Consideremos el suceso

$$A = \{w \,/\, X(w) \le x_2\} = \{w \,/\, X(w) \le x_1\} \cup \{w \,/\, x_1 < X(w) \le x_2\} = A_1 \cup A_2$$

Como $A_1 \cap A_2 = \emptyset,\ P(A) = P(A_1) + P(A_2)$, es decir

$$P(X \le x_2) = P(X \le x_1) + P(x_1 < X \le x_2) \ge P(X \le x_1)$$

y, por lo tanto,

$$F_X(x_2) \geq F_X(x_1)$$

iii) Recordemos que una función $g(x)$ es continua a derecha en *x* si $\lim_{h \to 0^+} g(x+h) = g(x)$.
Por lo tanto, la continuidad a derecha de $F_X(x)$ en todo *x* resulta de su definición: $F_X(x) = P(X \leq x)$.

iv) $\lim_{x \to \infty} F_X(x) = \lim_{x \to \infty} P(X \leq x) = \lim_{x \to \infty} P\{w / X(w) \leq x\} = P(S) = 1$

$$\lim_{x \to -\infty} F_X(x) = \lim_{x \to -\infty} P(X \leq x) = \lim_{x \to -\infty} P(w / X(w) \leq x) = P(\emptyset) = 0$$

v) $p_X(x) = P(X = x) = P(X \leq x) - P(X < x) = F_X(x) - F_X(x^-)$

Proposición: Sean a y b tales que $a \leq b$, entonces

$$P(a < X \leq b) = F_X(b) - F_X(a)$$
$$P(a \leq X \leq b) = F_X(b) - F_X(a^-)$$
$$P(a < X < b) = F_X(b^-) - F_X(a)$$
$$P(a \leq X < b) = F_X(b^-) - F_X(a^-)$$

Dem: Demostremos la primera igualdad

$$P(a < X \leq b) = P(X \in (a,b]) = P(X \in (-\infty, b]) - P(X \in (-\infty, a])$$
$$= P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$$

Ejercicio: Demostrar las siguientes 3 igualdades, usando por ejemplo que

$$P(a \leq X \leq b) = P(a < X \leq b) + P(X = a)$$

y aplicando la propiedad v) de las funciones de distribución acumuladas.

Ejemplo: Volviendo al ejemplo 1), y usando la función de distribución calculada antes, calculemos $P(1 \leq X \leq 2)$ y $P(X = 1)$.

$$P(1 \leq X \leq 2) = F_X(2) - F_X(1^-) = 1 - \frac{1}{4} = \frac{3}{4}$$
$$P(X = 1) = F_X(1) - F_X(1^-) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}.$$

Ejemplo: Un experimento tiene sólo dos resultados posibles, que denominaremos Éxito y Fracaso. El experimento se repite en forma independiente hasta que se obtiene el primer éxito. Sea *p* = P(Éxito), *0 < p < 1*, y definamos la v.a. *X* = "número de repeticiones hasta obtener el primer éxito". Como ya hemos visto, $R_X = N$.

Hallemos la función de probabilidad puntual de la v.a. *X.*

$$p_X(1) = p$$
$$p_X(2) = (1-p)\,p$$
$$p_X(3) = (1-p)^2\,p$$
....
$$p_X(k) = (1-p)^{k-1}\,p$$
....

Entonces,

$$p_X(k) = (1-p)^{k-1}\,p \qquad \forall k \in N.$$

Verifiquemos que en efecto esta función satisface las dos propiedades

$$p_X(x) \geq 0 \qquad \forall\, x$$

$$\sum_{x \in R_X} p_X(x) = 1$$

Dado que $0 < p < 1$, la primer propiedad obviamente se satisface. Respecto a la segunda,

$$\sum_{k=1}^{\infty} p_X(k) = \sum_{k=1}^{\infty} (1-p)^{k-1}\,p = p \sum_{j=0}^{\infty} (1-p)^j = p\,\frac{1}{1-(1-p)} = 1$$

donde hemos usado que la suma de la serie geométrica $\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$, si $|q| < 1$.

Hallemos la función de distribución acumulada de la v.a. *X.*

$x < 1 \qquad F_X(x) = 0$

$1 \le x < 2 \qquad F_X(x) = p$

$2 \le x < 3 \qquad F_X(x) = p + p(1-p)$

$3 \le x < 4 \qquad F_X(x) = p + p(1-p) + p(1-p)^2$

....

$$k \le x < k+1 \qquad F_X(x) = p\sum_{j=1}^{k}(1-p)^{j-1} = p\sum_{i=0}^{k-1}(1-p)^i = p\,\frac{1-(1-p)^k}{1-(1-p)} = 1-(1-p)^k$$

....

Hemos usado que la suma parcial de una serie geométrica es $\sum_{i=0}^{n} q^i = \frac{1-q^{n+1}}{1-q}$.

Recordemos que la función de distribución debe estar definida para todo $x \in \Re$, entonces

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1-(1-p)^{[x]} & \text{si } x \ge 1 \end{cases}$$

donde $[a]$ denota la parte entera de a.

Ejercicio: Verificar que esta función de distribución satisface las propiedades enunciadas antes.

Parámetro de una función de probabilidad: En el ejemplo anterior la probabilidad de Éxito la designamos *p* donde *0 < p < 1.* Variando este valor obtenemos diferentes funciones de probabilidad que constituyen lo que se denomina una *familia* de distribuciones. El valor *p* se denomina *parámetro* de la distribución.

En el caso del ejemplo, la familia obtenida se denomina Geométrica de parámetro *p* y diremos que *X ~ G(p)*. Por ejemplo, si el experimento hubiese consistido en arrojar un dado equilibrado hasta obtener el primer as, *X ~ G(1/6)* y si hubiese consistido en arrojar una moneda equilibrada hasta obtener la primera cara, *X ~ G(1/2).*

Esperanza o valor esperado de una v.a. discreta:

Una empresa proveedora de servicio de Televisión Satelital tiene 20000 clientes en cierta zona, cada uno de los cuáles puede optar por contratar de 1 a 5 paquetes de señales (el abono básico consiste en un solo paquete y cada uno de los otros paquetes incluye grupos de señales temáticas o premium). Supongamos que, entre los 20000 clientes, la distribución del número de paquetes *X* contratados es la siguiente:

x	**1**	**2**	**3**	**4**	**5**
número de clientes	7500	5500	3500	2000	1500
proporción	37.5%	27.5%	17.5%	10.0%	7.5%

Si interesa el número promedio de paquetes contratados, o sea el valor promedio de *X* en la población, deberíamos calcular:

$$\frac{1\cdot 7500 + 2\cdot 5500 + 3\cdot 3500 + 4\cdot 2000 + 5\cdot 1500}{20000} = \frac{44500}{20000} = 2.225$$

Observemos que, si no hubiésemos conocido los números de clientes que contratan cada número de paquetes ni el total de la población, sino sólo las proporciones de cada número (o su probabilidad) hubiésemos podido obtener el valor promedio, ya que dicho número puede escribirse en la forma:

$$1\cdot\frac{7500}{20000} + 2\cdot\frac{5500}{20000} + 3\cdot\frac{3500}{20000} + 4\cdot\frac{2000}{20000} + 5\cdot\frac{1500}{20000} =$$

$$= 1\cdot 0.375 + 2\cdot 0.275 + 3\cdot 0.175 + 4\cdot 0.10 + 5\cdot 0.075$$

Ésto motiva la siguiente definición.

<u>Definición</u>: Sea *X* una v.a. discreta que toma valores en R_X con función de probabilidad puntual $p_X(x)$, la **esperanza o valor esperado de *X*** se define como

$$E(X) = \mu_X = \sum_{x\in R_X} x\, p_X(x)$$

siempre que $\sum_{x\in R_X} |x|\, p_X(x) < \infty$. Si la serie de los valores absolutos diverge, la esperanza no puede definirse y decimos que no existe.

<u>Ejemplos</u>: 1) Sea *X*: "número de caras pares al arrojar dos veces un dado equilibrado". Como

x	0	1	2
$p_X(x)$	1/4	1/2	1/4

entonces, $E(X) = 0\frac{1}{4} + 1\frac{1}{2} + 2\frac{1}{4} = 1$.

2) Sea X una v.a. que toma sólo dos valores que designaremos 1 y 0 (Éxito y Fracaso) con la siguiente función de probabilidad puntual

x	1	0
$p_X(x)$	α	$1-\alpha$

siendo $0 < \alpha < 1$. Una v.a. de este tipo se dice que es una v.a. de tipo *Bernoulli* y su esperanza es:

$$E(X) = 1 \cdot \alpha + 0 \cdot (1-\alpha) = \alpha$$

3) Veamos un ejemplo en que no existe *E(X)*. Sea *X* una v.a. con la siguiente función de probabilidad puntual

$$p_X(x) = \begin{cases} \dfrac{6}{\pi^2}\dfrac{1}{x^2} & \text{si } x \in N \\ 0 & \text{en otro caso} \end{cases}$$

En primer lugar, observemos que $p_X(x)$ es una función de probabilidad puntual, ya que

$$\sum_{x=1}^{\infty} \frac{1}{x^2} = \frac{\pi^2}{6}$$

y, por lo tanto la suma de las probabilidades es 1. Calculemos la esperanza de *X*,

$$E(X) = \sum_{x=1}^{\infty} x \frac{6}{\pi^2} \frac{1}{x^2} = \frac{6}{\pi^2} \sum_{x=1}^{\infty} \frac{1}{x} = \infty$$

4) Consideremos nuevamente un experimento que tiene sólo dos resultados posibles y que se repite en forma independiente hasta que se obtiene el primer éxito. Si *p* = P(Éxito), $0 < p < 1$, y si definimos la v.a. *X* = "número de repeticiones hasta obtener el primer éxito", hemos demostrado que su función de probabilidad puntual está dada por

$$p_X(k) = (1-p)^{k-1} p \qquad \forall k \in N$$

Calculemos la esperanza de *X*.

$$E(X) = \sum_{k=1}^{\infty} k\; p(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = -p \sum_{k=1}^{\infty} \frac{\partial}{\partial p}(1-p)^k$$

Como la serie de potencias involucrada en la última igualdad es convergente, la derivada de la suma es la suma de las derivadas, entonces

$$E(X) = -p\,\frac{\partial}{\partial p}\left(\sum_{k=1}^{\infty}(1-p)^k\right) = -p\,\frac{\partial}{\partial p}\left(\frac{1}{1-(1-p)}-1\right) = -p\,\frac{\partial}{\partial p}\left(\frac{1}{p}-1\right) = -p\left(-\frac{1}{p^2}\right) = \frac{1}{p}.$$

y por lo tanto hemos demostrado que $E(X)=\dfrac{1}{p}$.

<u>Interpretación de la esperanza</u>: *E(X)* es el centro de gravedad de la función de probabilidad puntual. Es decir que si imaginamos que sobre cada valor posible de *X*, x_i, colocamos una masa $p_X(x_i)$, el punto de equilibrio del sistema es *E(X)*. En este sentido, podemos decir que *E(X)* es una medida del "centro" de la distribución.

Otra interpretación de *E(X)* está relacionada con un resultado que estudiaremos más adelante, denominado "ley de los grandes números". Imaginemos que se repite indefinidamente un experimento aleatorio y que en cada repetición nuestra v.a. *X* toma diferentes valores. Se ha demostrado que el promedio de los resultados obtenidos tiende a estabilizarse en un número que es *E(X)*, si es que ésta existe.

<u>Esperanza de una función de una v.a. discreta</u>: Volvamos al ejemplo considerado al comienzo del parágrafo dedicado a la esperanza. Sea la v.a. *X*: número de paquetes de programas contratado por un cliente seleccionado al azar y consideremos su función de probabilidad puntual:

x	1	2	3	4	5
$p_X(x)$	0.375	0.275	0.175	0.100	0.075

Supongamos que el costo del servicio (*Y*) es función del número de paquetes contratado, según la siguiente fórmula:

$$Y = 30\,(X+1)$$

¿Cuál es el valor esperado del costo pagado por cliente? Es decir, ¿cuál es *E(Y)*?.

A partir de la función de probabilidad puntual de *X*, podemos obtener la de función de probabilidad de *Y* ya que, por un lado R_Y = {60,90,120,150,180} y, por ejemplo, P(*Y*=120)=P(*X*=3)=0.175. Entonces,

y	60	90	120	150	180
$p_Y(y)$	0.375	0.275	0.175	0.100	0.075

y, $E(Y) = 60\cdot 0.375 + 90\cdot 0.275 + 120\cdot 0.175 + 150\cdot 0.10 + 180\cdot 0.075 = 96.75.$

Observemos que, $E(Y)=\sum_{x=1}^{5} h(x)p_X(x),$ siendo $h(x)=30(x+1).$

Proposición: Si X es discreta y toma valores $x_1, x_2, \ldots$, entonces $h(X)$ es discreta con valores $y_1, y_2, \ldots$, siendo $y_j = h(x_i)$ para al menos un valor de i.

Proposición: Si la v.a. X tiene función de probabilidad puntual $p_X(x)$ para todo $x \in R_X$, entonces la esperanza de cualquier función real $h(X)$, está dada por

$$E(h(X)) = \sum_{x\in R_X} h(x)\, p_X(x)$$

si la serie es absolutamente convergente, o sea si $\sum_{x\in R_X} |h(x)|\, p_X(x) < \infty$.

Dem: Sea $Y = h(X),$ entonces

$$E(Y)=\sum_j y_j\, p_Y(y_j) = \sum_j y_j \left[\sum_{i/h(x_i)=y_j} p_X(x_i)\right] = \sum_j \sum_{i/h(x_i)=y_j} y_j\, p_X(x_i) = \sum_i h(x_i)\, p_X(x_i).$$

Propiedades de la esperanza:

1) (*Linealidad*) Si a y b son constantes reales, $E(aX+b)=aE(X)+b$.

Dem: Sea $h(X)=aX+b,$ entonces

$$E(h(X))=E(aX+b)=\sum_{x\in R_X}(ax+b)\,p_X(x) = a\sum_{x\in R_X} x\, p_X(x) + b\sum_{x\in R_X} p_X(x) = aE(X)+b.$$

2) Si X es una v.a. tal que P(X=c)=1, entonces $E(X)=c$.

Dem: $E(X)=c\,p_X(c)=c.$

Varianza de una v.a. discreta:

Consideremos las siguientes funciones de probabilidad:

x	2	3	4
$p_X(x)$	1/3	1/3	1/3

y	1	2	3	4	5
$p_Y(y)$	1/12	5/12	2/12	1/12	3/12

z	3
$p_Z(z)$	1

Estas tres v.a. tienen la misma esperanza, sin embargo la forma de su distribución es muy diferente.

Ejercicio: Graficar las tres funciones de probabilidad puntual y verificar que *E(X)=E(Y)=E(Z)=3*.

Definiremos una medida de la variabilidad de una variable aleatoria alrededor de su esperanza.

Definición: Sea *X* una v.a. discreta con función de probabilidad puntual $p_X(x)$ y esperanza μ_X, la **varianza de *X***, que se denotará *V(X)*, σ_X^2 ó σ^2, es

$$V(X) = \sigma_X^2 = \sum_{x \in R_X} (x - \mu_X)^2 \, p_X(x) = E[(X - \mu_X)^2].$$

y el **desvío standard de *X***, es $\sigma_X = +\sqrt{V(X)}$.

Ejemplos: 1) Calculemos la varianza y el desvío standard de las tres v.a. que acabamos de presentar, cuya esperanza es igual a 3.

$$V(X) = \sigma_X^2 = (2-3)^2 \frac{1}{3} + (3-3)^2 \frac{1}{3} + (4-3)^2 \frac{1}{3} = \frac{2}{3}$$

$$V(Y) = \sigma_Y^2 = (1-3)^2 \frac{1}{12} + (2-3)^2 \frac{5}{12} + (3-3)^2 \frac{2}{12} + (4-3)^2 \frac{1}{12} + (5-3)^2 \frac{3}{12} = \frac{22}{12} = \frac{11}{6}$$

$$V(Z) = \sigma_Z^2 = (3-3)^2 \cdot 1 = 0$$

2) Consideremos *X*: “número de caras pares al arrojar dos veces un dado equilibrado” cuya función de probabilidad puntual es

x	0	1	2
$p_X(x)$	1/4	1/2	1/4

y su esperanza es $E(X) = 1$, entonces

$$V(X) = (0-1)^2 \frac{1}{4} + (1-1)^2 \frac{1}{2} + (2-1)^2 \frac{1}{4} = \frac{1}{2}.$$

3) Sea *X* una v.a. Bernoulli con función de probabilidad puntual

x	1	0
$p_X(x)$	α	$1-\alpha$

con $0 < \alpha < 1$. Recordemos que $E(X)=\alpha,$ entonces

$$V(X)=(1-\alpha)^2\alpha+(0-\alpha)^2(1-\alpha)=\alpha(1-\alpha)\left[(1-\alpha)+\alpha\right]=\alpha(1-\alpha).$$

Proposición: $V(X)=E(X^2)-\left(E(X)\right)^2$.

Dem:

$$V(X)=E\left((X-\mu_X)^2\right)=\sum_{x\in R_X}(x-\mu_X)^2\,p_X(x)=\sum_{x\in R_X}\left(x^2-2\mu_X x+\mu_X^2\right)p_X(x)=$$

$$=\sum_{x\in R_X}x^2\,p_X(x)-2\mu_X\sum_{x\in R_X}x\,p_X(x)+\mu_X^2\sum_{x\in R_X}p_X(x)=E(X^2)-2\mu_X E(X)+\mu_X^2=$$

$$=E(X^2)-2\mu_X^2+\mu_X^2=E(X^2)-\mu_X^2=E(X^2)-\left(E(X)\right)^2.$$

Ejemplo: Consideremos nuevamente un experimento que tiene sólo dos resultados posibles y que se repite en forma independiente hasta que se obtiene el primer éxito. Si *p* = P(Éxito), $0 < p < 1$, hemos definido la v.a. *X* = "número de repeticiones hasta obtener el primer éxito", cuya función de probabilidad puntual está dada por:

$$p_X(k)=(1-p)^{k-1}\,p \qquad \forall k\in N$$

Hemos demostrado que $E(X)=\dfrac{1}{p}$. Demostraremos ahora que $V(X)=\dfrac{1-p}{p^2}$.

Calculemos $E(X^2)$.

$$E(X^2)=\sum_{k=1}^{\infty}k^2\,p(1-p)^{k-1}=\sum_{k=1}^{\infty}\left[(k+1)k-k\right]p(1-p)^{k-1}=$$

$$=\sum_{k=1}^{\infty}(k+1)kp(1-p)^{k-1}-\sum_{k=1}^{\infty}k\,p(1-p)^{k-1}=\sum_{k=1}^{\infty}(k+1)kp(1-p)^{k-1}-E(X)=$$

$$=p\left[\sum_{k=1}^{\infty}(k+1)k(1-p)^{k-1}\right]-\frac{1}{p}=p\left[\sum_{k=1}^{\infty}\frac{\partial^2}{\partial p^2}(1-p)^{k+1}\right]-\frac{1}{p}=$$

$$= p \frac{\partial^2}{\partial p^2}\left[\sum_{k=1}^{\infty}(1-p)^{k+1}\right] - \frac{1}{p} = p \frac{\partial^2}{\partial p^2}\left[\sum_{j=2}^{\infty}(1-p)^{j}\right] - \frac{1}{p} =$$

$$= p \frac{\partial^2}{\partial p^2}\left[\frac{1}{1-(1-p)} - 1 - (1-p)\right] - \frac{1}{p} = p \frac{\partial^2}{\partial p^2}\left[\frac{1}{p} - 2 + p\right] - \frac{1}{p} =$$

$$= p \frac{\partial}{\partial p}\left[-\frac{1}{p^2} + 1\right] - \frac{1}{p} = p \frac{2}{p^3} - \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p}$$

Entonces,

$$V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p^2} - \frac{1}{p} = \frac{(1-p)}{p^2}$$

como queríamos demostrar.

Propiedades de la varianza y del desvío standard:

1) $V(aX+b) = a^2 V(X)$ y $\sigma_{aX+b} = |a|\,\sigma_X$.

Dem: Observemos que, en general,

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

Entonces,

$$V(aX+b) = \sum_{x \in R_X}\left(ax + b - E(aX+b)\right)^2 p_X(x) = \sum_{x \in R_X}\left(ax + b - aE(X) - b)\right)^2 p_X(x) =$$

$$= \sum_{x \in R_X}\left(ax - aE(X)\right)^2 p_X(x) = a^2 \sum_{x \in R_X}\left(x - E(X)\right)^2 p_X(x) = a^2 V(X)$$

y, por lo tanto, $\sigma_{aX+b} = |a|\,\sigma_X$.

En particular, observemos que $\sigma^2_{aX} = a^2\ \sigma^2_X$ y $\sigma^2_{X+b} = \sigma^2_X$, y por lo tanto un cambio de escala afecta la varianza pero una traslación no la afecta.

2) Si *X* es una v.a. tal que P(*X*=*c*) = 1, entonces *V(X) = 0*.

Dem: Ejercicio.