3.1: VARIABLE ALEATORIA

Comenzaremos por definir estrictamente qué es una variable estocástica o aleatoria:

> Según el Diccionario de la Real Academia Española
> VARIABLE : (Del latín *variabilis*) adj. Que varía o puede variar. ‖Inestable, inconstante y mudable.‖*Mat.* Magnitud que puede tener un valor cualquiera de los comprendidos en un conjunto.‖**estocástica**. Magnitud cuyos valores están determinados por las leyes de probabilidad, como los puntos resultantes de la tirada de un dado.
> ESTOCÁSTICA : (Del griego στοχαστικός, hábil en conjeturar) adj. Perteneciente o relativo al azar.

Introducción

Se desea realizar un control de calidad sobre cierto producto fabricado, llevando adelante el siguiente experimento estadístico: se toma una muestra aleatoria de tres piezas finalizadas y se clasifica cada una en dos posibles resultados: defectuosa y no defectuosa. A dichos eventos se los simboliza D y N, respectivamente. El espacio muestral resultante será:

Ω = {NNN, NND, NDN, NDD, DND, DDD}

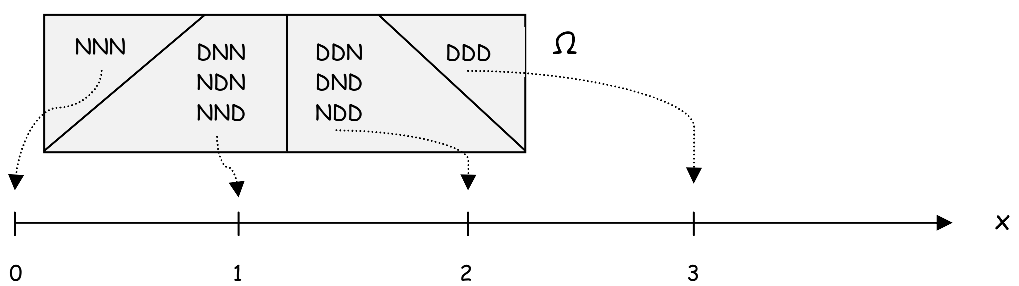
Luego definimos una cantidad X de la siguiente manera:

X: "Cantidad de piezas defectuosas que se encuentra en una muestra aleatoria de tamaño tres"

Esta cantidad puede tomar distintos valores en forma **aleatoria**; éstos se representan con x (equis minúscula) y son, para este experimento, X = {0, 1, 2, 3}.

Luego, a cada resultado del espacio muestral Ω podemos relacionarlo con uno de los valores que toma la cantidad X, es decir, con 0, 1, 2 y 3.

Gráficamente:



Se ha establecido así una relación que en términos matemáticos es una **función**, pues a cada resultado del espacio muestral se le ha asignado un resultado de la cantidad numérica *X*.
Este procedimiento puede generalizarse y establecer así un nuevo concepto, el de variable aleatoria.

Definición de variable aleatoria

Una variable aleatoria es una función que asigna a cada suceso del espacio muestral un número real. Dicho de otra forma, es una función definida en el espacio muestral que transforma sus resultados en puntos sobre la recta de los reales.
Simbólicamente:

$$\boxed{\begin{array}{rcl} X : \Omega & \rightarrow & \mathbb{R} \\ \omega & \mapsto & X(\omega) = x \end{array}}$$

El concepto de variable aleatoria proporciona un método para relacionar cualquier resultado con una variable cuantitativa.
En el ejemplo, los valores tomados por la variable *X* son 0, 1, 2 y 3 y se interpretan como "cero piezas defectuosas en una muestra aleatoria de tamaño tres", "una pieza defectuosa en una muestra aleatoria de tamaño tres" e interpretaciones análogas para los valores restantes.

Otras importantes definiciones

- Se dice que una variable aleatoria es **discreta** si la cantidad de valores que puede tomar es enumerable o contable (finita o infinita), y si estos pueden arreglarse en una secuencia que corresponde con los números enteros.

- Se dice que una variable aleatoria es **continua** si los valores que puede tomar corresponden a un intervalo de la recta de los reales.

Veamos algunas aplicaciones: En el ejemplo de las piezas defectuosas definimos P(*X*=*x*) como la probabilidad de que la variable aleatoria tome exactamente alguno de los valores 0, 1, 2 o 3. En contexto, para *x* = 1, la P(*X*=1) se interpretaría como la probabilidad de que la cantidad de piezas defectuosas en una muestra aleatoria de tamaño tres sea exactamente una. Dado que hay tres posibles resultados en el espacio muestral en los que aparece una pieza defectuosa en la muestra de tres {NND, NDN, DNN} de entre un total de ocho resultados que conforman el espacio muestral, mediante la definición clásica de probabilidad podemos escribir P(*X* = 1) = 3/8[1].
En la Tabla 1, que está a continuación, se muestran los valores de probabilidad para las cantidades de piezas defectuosas que pueden extraerse.

---

[1] Aplicando sucesivamente la regla de la suma para eventos mutuamente excluyentes y la regla de la multiplicación para eventos independientes, realizando cada extracción de una pieza cualquiera con reemplazo, y considerando a N y D equiprobables con p=0,5 se llega al mismo resultado:

$$P\left[(NND)\cup(NDN)\cup(DNN)\right] = P(N\cap N\cap D) + P(N\cap D\cap N) + P(D\cap N\cap N) = P(N)P(N)P(D) + P(N)P(D)P(N) + P(N)P(N)P(D) =$$

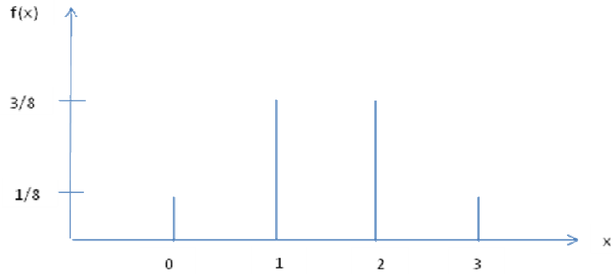
$$= 3\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} = \frac{3}{8}$$

De esta forma, al considerar los valores de una variable aleatoria, es posible desarrollar una función matemática que asigne a cada realización *x* de la variable, una probabilidad. Esta función recibe el nombre de **función distribución de probabilidad** de la variable aleatoria *X* o también **función masa de probabilidad** de la variable aleatoria X, cuyo símbolo es *f*(*x*). La distribución de probabilidad para el ejemplo se muestra en la Figura 1 que está a continuación de la tabla.

Tabla 1

x	P(X = x)	P(X ≤ x)
0	P(X=0) = 1/8	P(X ≤ 0) = 1/8
1	P(X=1) = 3/8	P(X ≤ 1) = 4/8
2	P(X=2) = 3/8	P(X ≤ 2) = 7/8
3	P(X=3) = 1/8	P(X ≤ 3) = 8/8

Figura 1



- Para toda variable aleatoria discreta X se llamará a *f*(*x*) = P(*X*=*x*) **función distribución de probabilidad** (o **función masa de probabilidad**) si satisface las siguientes propiedades[2]:

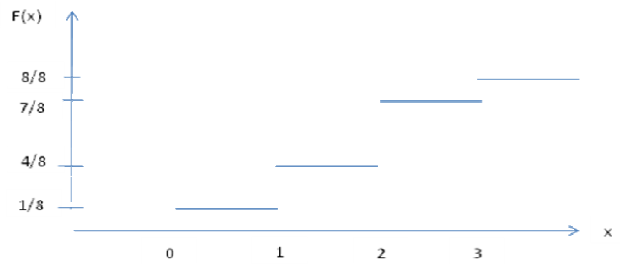
  1. $f(x) \geq 0,$ para todo valor x de X
  2. $\sum_{x} f(x) = 1$

- Una función de distribución de probabilidad también puede definirse teniendo en cuenta los valores de la variable **hasta** uno en particular. En la tabla precedente puede verse que se ha calculado P(*X* ≤ *x*), la cual se interpreta, en el contexto dado, como la probabilidad de obtener **hasta** *x* piezas defectuosas cuando se extraen aleatoriamente de una muestra de tamaño tres. A esta función se la conoce como **función distribución de probabilidad acumulada** de la variable *X*, cuyo símbolo es *F*(*x*).

  Para toda variable aleatoria discreta *X*, se llamará **función distribución acumulada de probabilidad** a la probabilidad de que la variable aleatoria *X* tome valores hasta cierto valor *x* incluido dentro de su rango de definición. En símbolos: $F(x) = P(X \leq x) = \sum_{x_i} f(x_i); \quad (x_i \leq x)$

[2] Estas propiedades son análogas a los axiomas de probabilidad P(A) ≥ 0, (∀A) (A∈Ω) y P(Ω)=1

Gráficamente:



Notamos que:

- $F(1) = P(X \le 1) = f(0) + f(1)$
- $F(2) = P(X \le 2) = f(0) + f(1) + f(2)$ siendo $f(2) = P(X = 2) = F(2) - F(1)$
- $P(X \ge 2) = f(2) + f(3)$

Recordando que $\sum_{x=0}^{3} f(x) = F(3) = 1$

$$\rightarrow P(X \ge 2) = \sum_{x=0}^{3} f(x) - \sum_{x=0}^{1} f(x) = F(3) - F(1) = 1 - F(1)$$

$$\rightarrow P(X > 2) = \sum_{x=0}^{3} f(x) - \sum_{x=0}^{2} f(x) = F(3) - F(2) = 1 - F(2)$$

Podemos generalizar los resultados anteriores de la siguiente manera:

1. $P(X = x_i) = F(x_i) - F(x_{i-1})$
2. $P(X \ge x_i) = 1 - F(x_{i-1})$
3. $P(X > x_i) = 1 - F(x_i)$
4. $P(x_i \le X \le x_j) = F(x_j) - F(x_{i-1})$ $\quad (\forall x_j)(x_j \ge x_i)$
5. $P(x_i < X < x_j) = F(x_{j-1}) - F(x_{i+1})$ $\quad (\forall x_j)(x_j > x_i)$

Si *X* es una variable aleatoria discreta, la función distribución de probabilidad, *f*(*x*), es la probabilidad de que la variable tome algún valor dentro de su rango de definición, siendo estos valores cantidades discretas, es decir, numerables. Supongamos que *X* toma valores en un intervalo real [a ; b], es decir, a ≤ X ≤ b (donde [a ; b] ∈ IR). Entonces X es una variable aleatoria continua; por lo tanto toma infinitos valores posibles en el intervalo [a ; b] dado que *IR* es un conjunto denso. Vemos ahora que P(*X* = x_i) = 0 necesariamente, pues la probabilidad de que la variable tome **exactamente** el valor x_i y no cualquier otro se considera nula (*i* es un índice que recorre todos los valores de variable); dicho de otra manera, **observar** exactamente un valor x_i de la variable y no cualquier otro infinitesimalmente cercano no es posible en términos probabilísticos[3]. Como P(*X* = x_i) = 0 **siempre**, *f*(*x*)

[3] Esto no debe interpretarse de ninguna manera como que el valor x_i no existe; la variable es continua, lo que quiere decir que no tiene saltos o discontinuidades en su dominio.

ya no representa la probabilidad de que la variable X tome un valor x cualquiera en su dominio. Si X es una variable aleatoria continua, $f(x)$ ahora proporciona un **medio** para determinar la probabilidad en un intervalo $a \leq X \leq b$ pero **no es** de ninguna manera **una probabilidad** en sí misma como sí lo es en el caso de variables aleatorias discretas.

Para entender cómo se utiliza $f(x)$ para obtener probabilidades volvamos al modelo discreto, para un caso cualquiera en el que X toma algún valor x_i, donde i = 0, 1, 2, 3, …, 10.

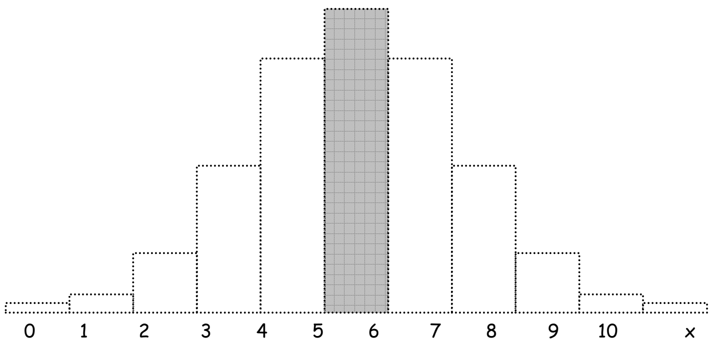
Puede observarse en la Figura 2 que si tomamos un rectángulo de base unitaria centrado en cualquiera de los valores x, por ejemplo en 5, la base b estará acotada en el intervalo $4{,}5 \leq x \leq 5{,}5$ y la altura h será justamente el valor de $f(x)$ en x=5; entonces:

El área de dicho rectángulo será $A = b \cdot h = \Delta x \cdot f(5) = 1 \cdot f(5) = P(X = 5)$.

En general, para cualquier rectángulo centrado en x_i, $A_i = f(x_i) = P(X = x_i)$.

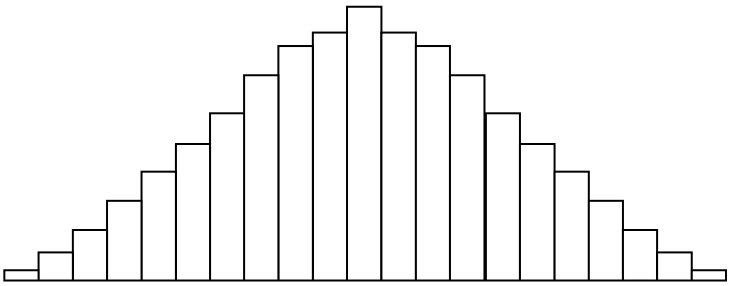
Este importante resultado nos dice que **una probabilidad puede interpretarse numéricamente como un área**, en este caso, el área del rectángulo de base unitaria centrado en el valor x_i de la variable y cuya altura es exactamente $f(x_i)$. Así, $A_T = \sum A_i = \sum f(x_i) = 1$, de acuerdo con la segunda propiedad de las distribuciones de probabilidad de variables aleatorias discretas.

Figura 2



Si ahora mantenemos fijos los extremos pero aumentamos la cantidad de rectángulos con los que subdividimos el intervalo (y suponiendo, únicamente con fines ilustrativos, que la variable puede tomar más valores dentro de ese rango a intervalos regulares) obtendríamos la siguiente figura:

Figura 3



Este esquema puede continuar, con cada vez más valores de la variable *X*. Cuando el número observado de valores de la variable sea muy grande y la amplitud de los intervalos sea muy pequeña (o, equivalentemente, el área de los rectángulos individuales cada vez menor) la frecuencia relativa, es decir, las probabilidades de ocurrencia de x_i, aparecerá como una curva suave. Con base en la Figura 3 puede especularse que la curva límite para el ejemplo es la que se muestra en la Figura 4

- La función $f(x)$, cuya gráfica se obtiene para un número muy grande de observaciones y para una amplitud de intervalo muy pequeña es la **función densidad de probabilidad** de una variable aleatoria continua X.
  Como ya se vio, el área total de los rectángulos es $A_T = \sum A_i = \sum f(x_i)\Delta x = 1$; podemos escribir ahora la suma de Riemann $\sum_{i=1}^{n} f(x_i)\cdot\Delta x$, y tomar el límite para una cantidad de rectángulos tan grande como se quiera con bases tan pequeñas como se quiera, lo que es el área total bajo la curva y es la integral definida entre los dos valores extremos considerados de la variable aleatoria X.
  En símbolos, $\lim_{n\to\infty}\sum_{i=1}^{n} f(x_i)\cdot\Delta x = \int_a^b f(x)dx = A(x)$. Esto puede verse en la Figura 5.

Figura 4

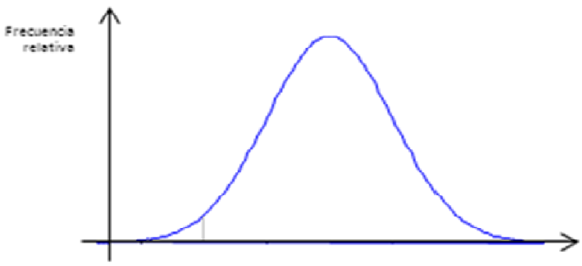
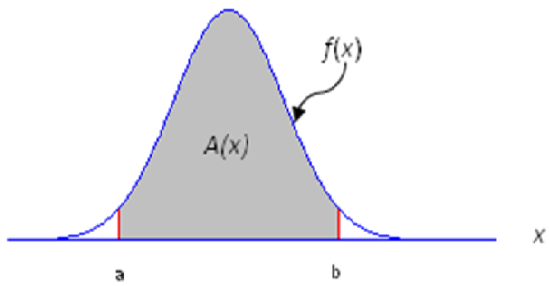


Figura 5



- Así, la **función densidad de probabilidad de una variable aleatoria *X* continua** se define de la siguiente manera:
  Si existe una función $f(x)$ tal que cumpla las propiedades:
  1. $f(x) \geq 0, \ -\infty < x < \infty$
  2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

  entonces $f(x)$ es la **función densidad de probabilidad** de dicha variable aleatoria continua. La propiedad 3 es consecuencia de la 2[4].
  Como consecuencia de la propiedad 2
  $P(a \leq X \leq b) = P(a < X < b) = \int_a^b f(x)\,dx = A(x)$
  para cualesquiera *a* y *b* en el dominio de *X*

[4] Nótese que las propiedades 1 y 2 son análogas a las de las variables discretas. Sin embargo no significan lo mismo, dado que $f(x)$ ya no es una probabilidad, como sabemos, sino que ahora es solamente la función matemática que modela la distribución de frecuencias de la variable *X*. En este caso $f(x) \geq 0$ significa que la función es siempre positiva; es decir, que siempre está por encima del eje x.

- Al igual que en el caso de una variable aleatoria discreta, la **función distribución de probabilidad acumulada** de una variable aleatoria continua X es la probabilidad de que la variable tome algún valor menor o igual a algún x específico.

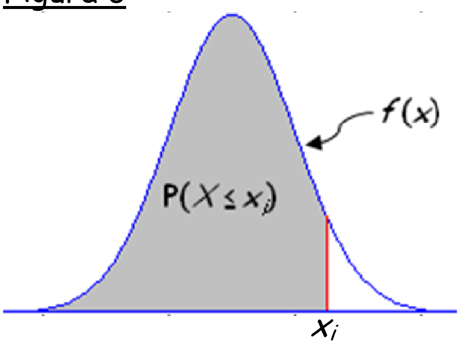
  En símbolos: $P(X \le x) = \int_{-\infty}^{x} f(t)\,dt = F(x)$

  Se cumple para $F(x)$ que

  1. $F(-\infty) = 0$
  2. $F(+\infty) = 1$
  3. $P(a < X < b) = F(b) - F(a)$
  4. $\dfrac{dF(x)}{dx} = f(x)$

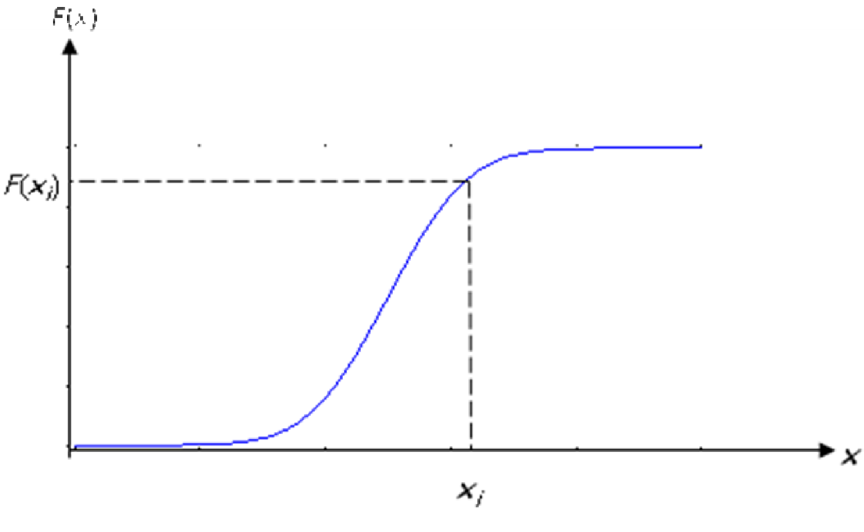
  La propiedad 4 es una consecuencia del Teorema Fundamental del Cálculo mientras que la 3 es la Regla de Barrow. Puede verse en la Figura 6 a la distribución acumulada como el área bajo la función densidad hasta el valor x_i considerado.

Figura 6



La representación gráfica de F(x) es, en cambio, una curva creciente (su aspecto recuerda a una ojiva) y en la cual, cabe destacar, no se representan áreas como probabilidades sino que éstas se encuentran en el eje de las frecuencias acumuladas para las correspondientes abscisas, tal como muestra la Figura 7.

Figura 7



Valor esperado de una variable aleatoria

El **valor esperado** (o **esperanza matemática**) de una variable aleatoria es un concepto muy importante en el estudio de las distribuciones de probabilidad. Se entiende como el *valor promedio que toma una variable aleatoria después de un número grande de experimentos*[5].

En el ejemplo de las tres piezas extraídas al azar en búsqueda de defectuosas vemos que, de acuerdo a la Tabla 1, esperamos obtener ninguna defectuosa en una de cada ocho extracciones, una defectuosa en tres de cada ocho extracciones, dos defectuosas en tres de cada ocho extracciones y las tres defectuosas en una de cada ocho extracciones. Si ponderamos estos valores esperados para cada resultado con sus respectivas probabilidades y luego los sumamos obtenemos:

$$0\times\frac{1}{8}+1\times\frac{3}{8}+2\times\frac{3}{8}+3\times\frac{1}{8}=\frac{3}{2}=1{,}5\ \text{(defectuosas)}$$

Nótese que, tratándose X de la cantidad de piezas defectuosas en una muestra aleatoria de tres, el valor 1,5 es la cantidad de piezas defectuosas que se espera obtener si se realiza la extracción de tres piezas aleatoriamente un gran número de veces. Es decir, este valor obtenido, el valor esperado o promedio, es un valor teórico o límite. Esto es así debido a que las respectivas probabilidades para los valores de x también son las teóricas, pues no esperamos que si realizamos ocho veces el experimento exactamente una sola vez aparezcan cero defectuosas, tres veces una defectuosa, tres veces dos defectuosas y una vez las tres defectuosas; sabemos que esto ocurrirá solamente cuando hayamos repetido una gran cantidad de veces el experimento de extraer las tres piezas al azar[6].

Del ejemplo se generaliza la siguiente definición:

- El valor esperado (o esperanza matemática) de una variable aleatoria X está dado por:

  $$\mu_X = E(X) = \sum_x x\,.\,f(x) \qquad \text{si X es discreta} \qquad [\,1\,]$$

  $$\mu_X = E(X) = \int_{-\infty}^{\infty} x\,.\,f(x)\,dx \qquad \text{si X es continua} \qquad [\,2\,]$$

  Donde $f(x)$ es la función masa de probabilidad y la función densidad de probabilidad, que corresponde a cada caso.

  Se interpreta al valor esperado como "**el valor promedio que se espera que resulte cuando se realiza una gran cantidad de veces el experimento**"[7]. En el ejemplo el valor 1,5 no es un valor que la variable alguna vez pueda tomar,

---

[5] La esperanza tiene su origen en los juegos de azar, dado que los apostadores deseaban saber cuál era su esperanza de ganar repetidamente un juego. Así, el valor esperado representa la cantidad de dinero promedio que el jugador está dispuesto a ganar o perder después de un número muy grande de apuestas. En Teoría de las Decisiones juega un rol fundamental al permitir proyectar las ganancias (o pérdidas) esperadas cuando se toman decisiones bajo incertidumbre.

[6] Recordemos que los valores de probabilidad obtenidos mediante la definición frecuencial tienden a los predichos por la definición clásica cuando el número de ensayos es muy grande.

[7] Por otro lado, el valor esperado podría no existir si la correspondiente sumatoria o integral no converge a un valor finito.

dado que se trata de una variable discreta; sin embargo se está hablando del valor medio (o promedio), el cual siempre es un valor comprendido en el campo de los reales y que refleja la cantidad de defectuosos que se espera que aparezcan a lo largo del tiempo cuando se haya realizado una gran cantidad de veces las extracciones de tres piezas aleatoriamente.

Es fácil darse cuenta, observando el caso discreto, por qué se entiende al valor esperado como el valor promedio de la variable. La expresión $\sum xf(x)$ no es otra cosa que un promedio ponderado, pues se está sumando cada valor de la variable multiplicado por su probabilidad (la cual puede entenderse como el peso relativo si la relacionamos con la frecuencia de aparición).

La esperanza de una variable aleatoria *X* no es una función de *X* sino un número fijo, y es una característica de la distribución de probabilidad de *X*.

Se define el valor esperado de una función de una variable aleatoria como una generalización del concepto de valor esperado.

- Si $g(X)$ es cualquier función de una variable aleatoria X, entonces el valor esperado de esta función de X es

$$E[g(X)] = \sum_x g(x)f(x) \qquad \text{si X es discreta} \qquad [3]$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx \qquad \text{si X es continua} \qquad [4]$$

Así, por ejemplo, si X es una variable aleatoria discreta y $g(x)=x^2$, entonces:

$$E[g(X)] = E[X^2] = \sum_x x^2 f(x)$$

<u>Ejemplo</u>: El número de autos atendidos en un lavadero un sábado por la tarde es una variable aleatoria discreta cuya función masa de probabilidad es:

x	4	5	6	7	8	9
f(x)	1/12	1/12	1/4	1/4	1/6	1/6

La cantidad, en pesos, de dinero que reciben los empleados está fijada de acuerdo a la cantidad de autos que ingresan mediante la relación $g(x)=20x-10$.

¿Cuánto dinero esperan ganar los empleados una tarde cualquiera de sábado?

Podemos agregar una fila más a la tabla con los valores de $g(X)$ correspondientes. Así,

x	4	5	6	7	8	9
f(x)	1/12	1/12	1/4	1/4	1/6	1/6
g(X)	70	90	110	130	150	170

Entonces,

$E[g(X)]=E(20x-10)=$

$$= \sum_{x=4}^{9} (20x-10)f(x) = 70\times\frac{1}{12} + 90\times\frac{1}{12} + 110\times\frac{1}{4} + 130\times\frac{1}{4} + 150\times\frac{1}{6} + 170\times\frac{1}{6} \approx \$126{,}67$$

¿Cómo se interpreta este resultado? La interpretación formal es "se espera que, en promedio, los empleados ganen $126,67 una tarde cualquiera de sábado". ¿Pero qué nos quiere decir esto, realmente? Por supuesto que no es realista pensar que

cada sábado ganarán esa cantidad exactamente; en realidad depende de cuántos automóviles entren al lavadero, y las cantidades 4, 5, 6, 7, 8 y 9 son **verdaderamente aleatorias**, siendo las probabilidades correspondientes las frecuencias con que estos resultados han aparecido a lo largo de un tiempo determinado en el que se ha hecho el estudio. Así, algunas tardes de sábado ganarán más que eso y otras menos, pero si se considera un tiempo suficientemente largo (bajo las mismas condiciones que llevaron a establecer $f(x)$) entonces \$126,67 es el promedio de las ganancias concretas que tendrán durante todos los sábados por la tarde que trabajen.
Puede que este resultado sea un poco decepcionante, pero en realidad tiene mucha utilidad. Por un lado, como ya se dijo, en Teoría de Decisiones el concepto juega un rol fundamental al permitir proyectar las ganancias (o pérdidas) esperadas en grandes volúmenes de operaciones cuando se toman decisiones bajo incertidumbre. Por otro lado, en el estudio de modelos probabilísticos, el valor esperado es verdaderamente la media aritmética de las distribuciones de probabilidad (el valor promedio para todos los valores de la variable) y tiene aplicaciones fundamentales en, por ejemplo, el control de calidad o en el campo de la inferencia estadística.

<u>Ejemplo</u>: La proporción de personas que responde a una encuesta viene dada por la función densidad

$$f(x) = \begin{cases} \dfrac{2(x+2)}{5} & 0 < x < 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Además, la proporción de personas que responden en tiempo y forma es una función de X tal que $g(X) = (2/3)X$. ¿Qué proporción de personas se espera que respondan en tiempo y forma la encuesta?
Empleando la expresión [4]

$$E[g(X)] = \int_0^1 \frac{2x}{3}\frac{2(x+2)}{5}dx = \frac{4}{15}\int_0^1 (x^2+x)dx = \frac{4}{15}\left[\frac{x^3}{3}+\frac{x^2}{2}\right]_0^1 = \frac{2}{9}$$

Un ***<u>caso particular importante</u>*** es cuando $\boldsymbol{g(X)=(X-\mu)^2}$, donde $\mu = E(X)$. De acuerdo a la definición, si X es una variable aleatoria discreta, entonces:

$$E[g(X)] = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 f(x)$$

Puede verse que, como resultado de aplicar la definición, se está sumando las diferencias de cada valor de la variable aleatoria X con respecto a la media de la misma elevadas al cuadrado y promediadas por su frecuencia de aparición (es decir, ponderada por su probabilidad). Así, se ha hecho un promedio de las diferencias cuadráticas de la variable con respecto a su propia media. Esto no es otra cosa que la varianza de la variable aleatoria. Se introduce, entonces, la siguiente definición.

Varianza de una variable aleatoria

- Si *X* es una variable aleatoria, entonces su **varianza** está dada por

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 f(x) \qquad \text{si X es discreta} \qquad [5]$$

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \qquad \text{si X es continua} \qquad [6]$$

Para el muestreo de las piezas defectuosas, tenemos:

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = \sum_{x=0}^{3} (x-\mu)^2 f(x) =$$

$$= (0\text{-}1{,}5)^2 \times \frac{1}{8} + (1-1{,}5)^2 \times \frac{3}{8} + (2\text{-}1{,}5)^2 \times \frac{3}{8} + (3-1{,}5)^2 \times \frac{1}{8} = 0{,}75 \text{ (defectuosas)}^2$$

(Recordemos que, por tratarse de una varianza, las unidades están elevadas al cuadrado y que, en este caso, son piezas defectuosas. Como siempre, se puede encontrar el desvío estándar como la raíz cuadrada de la varianza y expresar la dispersión de la variable en estudio en términos de las unidades originales)

<u>Fórmula alternativa</u>

Puede verse que si la cantidad de valores *x* que toma la variable fuera importante, el cálculo de la varianza se volvería muy tedioso.

Existe una fórmula alternativa a [5] y [6], la cual es:

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = E(X^2) - [E(X)]^2 \qquad [7]$$

<u>Demostración</u>

$$\sigma_X^2 = E\left[(X-\mu)^2\right] = E\left[(X^2 - 2X\mu + \mu^2)\right] \qquad \text{desarrollando el cuadrado del binomio}$$

$$= \sum_x (x^2 - 2x\mu + \mu^2) f(x) \qquad \text{aplicando la definición [6]}$$

$$= \sum_x x^2 f(x) - \sum_x 2\mu x f(x) + \sum_x \mu^2 f(x) \qquad \text{aplicando propiedad distributiva}$$

$$= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x) \qquad \text{ya que 2 y } \mu \text{ son constantes}$$

$$= \sum_x x^2 f(x) - 2\mu\mu + \mu^2 \qquad \text{ya que } \sum_x x f(x) = E(X) = \mu \text{ por [1] y } \sum_x f(x) = 1 \text{ por propiedad de } f(x)$$

$$= \sum_x x^2 f(x) - 2\mu^2 + \mu^2$$

$$= \sum_x x^2 f(x) - \mu^2 \qquad \text{y como } x^2 \text{ es una función } g(x) \text{ de X}$$

$$\Rightarrow \sigma_X^2 = E(X^2) - [E(X)]^2$$

Variable aleatoria

La expresión [7] se desarrolló a partir del caso discreto, pero es válida también para el caso continuo.

Así, para el caso de las piezas defectuosas,

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

$$[E(X)]^2 = \left(\frac{3}{2}\right)^2 = \frac{9}{4}$$

Entonces, $\sigma_X^2 = E(X^2) - [E(X)]^2 = 3 - \frac{9}{4} = \frac{3}{4} = 0{,}75$

Que es el mismo resultado al que se llegó aplicando [6]

¿Cuál es la varianza para la proporción de personas que responden a la encuesta en el ejemplo?

$$E(X^2) = \int_0^1 x^2 dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

$$E(X) = \int_0^1 x dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2} \Rightarrow [E(X)]^2 = \left[\frac{1}{2}\right]^2 = \frac{1}{4}$$

$$\sigma_X^2 = E(X^2) - [E(X)]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \approx 0{,}083$$

$$DE(X) = \sqrt{\sigma_X^2} = \sigma_X \approx 0{,}29$$

Propiedades del valor esperado y de la varianza

- $E(c) = c$. **"El valor esperado de una constante c es la misma constante"**

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

- $E(aX + b) = aE(X) + b$. **"El valor esperado de una cantidad *aX+b*, donde *a* y *b* son constantes, es el producto de *a* por el valor esperado de *X*, más *b*"**

  $$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)f(x)dx = \int_{-\infty}^{\infty} axf(x)dx + \int_{-\infty}^{\infty} bf(x)dx =$$

  $$= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx = aE(X) + b$$

- $E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$ **"El valor esperado de la suma (o resta) de dos funciones de *X* es la suma (o resta) de los valores esperados de las funciones de *X*"**

  $$E[g(X) \pm h(X)] = \int_{-\infty}^{\infty} [g(x) \pm h(x)]f(x)dx = \int_{-\infty}^{\infty} g(x)f(x)dx \pm \int_{-\infty}^{\infty} h(x)f(x)dx =$$

  $$= E[g(X)] \pm E[h(X)]$$

- $\sigma^2_{aX+b} = a^2\sigma^2_X$ **"Si a y b son constantes, la varianza de la cantidad *aX+b* es la varianza de *X* multiplicada por el cuadrado de la constante a"**

  $$\sigma^2_{aX+b} = \left\{E\left[(aX+b)-\mu_{aX+b}\right]^2\right\} = E\left[(aX+b)-a\mu-b\right]^2 = E\left[(aX-a\mu)^2\right] =$$
  $$= E\left[a^2(X-\mu)^2\right] = a^2E\left[(X-\mu)^2\right] = a^2\sigma^2_X$$

  Como caso particular, si $a = 1 \Rightarrow \sigma^2_{X+b} = \sigma^2_X$. Puede verse que la varianza de *aX+b* no depende de la constante *b*, pues no aparece en el miembro de la derecha de la igualdad. Esto se debe a que la dispersión de una variable cualquiera no se ve afectada por la suma de una constante a la variable. No obstante, esta adición sí afecta al valor esperado.

Teorema de Chebyschev

La varianza de una variable aleatoria dice algo acerca de la variabilidad de las observaciones alrededor de la media. Si una variable aleatoria tiene una varianza (o una desviación estándar) pequeña, se espera que la mayor parte de los valores se agrupen cerca de la media alrededor de ella. Así, la probabilidad de que la variable aleatoria tome valores dentro de un intervalo alrededor de la media es mayor que para una variable aleatoria similar con mayor varianza.
Como el área bajo una curva de distribución (o en un histograma de probabilidad, si se trata de una variable aleatoria discreta) suma 1, el área entre dos valores cualesquiera es la probabilidad de que la variable aleatoria tome un valor entre estos dos extremos, tal como dice la propiedad (3) de las distribuciones de probabilidad para variables continuas.
El siguiente teorema proporciona una herramienta para calcular las probabilidades de una variable aleatoria *X*, discreta o continua, al tomar valores en un intervalo, en casos de incertidumbre, esto es, cuando se desconoce su distribución de probabilidad.

Sea *X* una variable aleatoria cualquiera, la probabilidad de que tome algún valor dentro de un intervalo simétrico alrededor de la media dentro de k desviaciones estándar es <u>al menos</u> 1 – 1/*k*^{2}.

En símbolos: $P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2} \qquad k \in \mathbb{R}^+$

Este teorema da una estimación **conservadora** de la probabilidad buscada; esto quiere decir que la relación de "mayor o igual" asegura un límite inferior pero nunca el que verdaderamente corresponde, pues a éste solamente se llegaría si se conociera la distribución de probabilidad de la variable.
La utilidad del teorema radica en, precisamente, obtener información (aunque no exacta) con respecto a probabilidades referidas a la variable aleatoria *X* en condiciones de incertidumbre.

Ejemplo: El ingreso monetario en cierta población se puede modelar mediante una variable aleatoria continua, cuya función densidad de probabilidad no se ha estudiado. ¿Bajo estas condiciones, si se elije una persona al azar de esta población, cuál es la probabilidad de que el ingreso esté dentro de los dos desvíos estándar con respecto al ingreso medio?

X: cantidad (en unidades monetarias) ganada por las personas de la población considerada.

De acuerdo al enunciado, lo que se pide es la probabilidad para X en un intervalo $|X-\mu|<k\sigma$ donde $k=2$. Entonces: $|X-\mu|<2\sigma$

$-2\sigma < X-\mu < 2\sigma$

$\mu-2\sigma < X < \mu+2\sigma$, por lo tanto

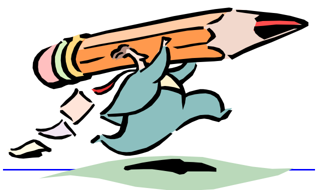
$P(\mu-2\sigma < X < \mu+2\sigma)$

Como se desconoce $f(x)$, no es posible encontrar exactamente la probabilidad pedida. Sin embargo, de acuerdo al teorema de Chebyschev

$$P(\mu-2\sigma < X < \mu+2\sigma) \geq 1-\frac{1}{2^2}=\frac{3}{4}$$

Es decir, que la probabilidad de que el ingreso de una persona elegida al azar esté dentro de los dos desviaciones estándar con respecto al ingreso medio es de, **al menos**, 0,75.

La probabilidad exacta sólo puede encontrarse si se conoce la función densidad de probabilidad pero, ante la falta de información, este resultado es la diferencia con no tener nada.



A trabajar solos...

Aunque no tanto porque al final encontrará los ejercicios resueltos.

1. Clasifique las siguientes variables aleatorias como discretas o continuas:
   A: Producción diaria de motores en una empresa automotriz.
   B: Tiempo de espera hasta acceder a un servidor.
   C: Venta, en dólares, de cada vendedor de un comercio.
   D: Vida útil de un rodamiento.
   E: Cantidad de unidades defectuosas en una línea de producción.
   F: Nivel de contaminación ambiental.
   G: Cantidad de intentos hasta acertar en un blanco.
   H: Sueldo de los empleados de cierta industria.

2. El grado con el cual una compañía se involucra con el desarrollo de nuevos productos depende, en gran medida, de sus competidores. Un estudio fue realizado sobre las compañías que utilizan tecnología avanzada, para analizar el número de nuevos productos que se lanzan al mercado anualmente. Los cálculos, en función de años anteriores, están dados en la siguiente tabla:

X	3	4	5	6	7	8	9
f(x)	0,08	0,14	0,22	0,30	0,14	0,08	0,04

a) Grafique la función de cuantía y la función de probabilidad acumulada.

b) Calcule el valor esperado, la varianza y la desviación estándar.

c) Encuentre la proporción de compañías de tecnología avanzada para las cuales la variable aleatoria iguala o excede a $\mu + 2\sigma$.

3. Una caja contiene tres bolillas numeradas del 1 al 3, primero se extrae una bolilla y luego se lanza una moneda legal tantas veces como indique la bolilla. Calcule f(x), F(x) y grafique, siendo X: 'número de caras obtenidas'.

4. Se lanza una moneda insesgada hasta que salga una cara o cinco cruces. Encuentre el número esperado de lanzamientos de la moneda.

5. Se sabe por experiencia que la demanda diaria de un producto perecedero es como se muestra en la siguiente tabla:

X	3	4	5	6	7	8	9
f(x)	0,05	0,12	0,20	0,24	0,17	0,14	0,08

a) Grafique la función de cuantía y la función de probabilidad acumulada.

b) Si cada artículo tiene un costo total de \$ 45 con un precio de venta de \$ 75, ¿cuál es la utilidad esperada?

6. Un jugador lanza dos monedas no cargadas. Gana \$2 si salen 2 caras y \$1 si sale 1 cara. Por otra parte, pierde \$3 si no sale ninguna cara. Determine el valor esperado del juego y si éste es favorable al jugador. Si el juego debe ser justo, ¿cuántos debería perder si no sale ninguna cara?

7. Se lanza una moneda al aire 3 veces. Si se obtienen al menos 2 caras se permite tirar un dado y recibir en dólares el número de puntos obtenido. ¿Qué es lo que puede esperar ganar en este juego en un intento?

8. Dada la variable aleatoria continua X con función de densidad $f(x)=k.x^2.(1-x)$ para $0 \le x \le 1$ y 0 para otro caso.

a) Determine la constante k.

b) Halle F(x).

9. Sea la variable aleatoria continua X: "duración en horas de un circuito electrónico" con función de densidad:

$$f(x) = \begin{cases} 500 / x^2 & x \ge 500 \\ 0 & x < 500 \end{cases}$$

a) Pruebe que f(x) es función de densidad.

b) Halle F(x).

c) ¿Cuál es la probabilidad de que un circuito dure entre 1500 y 2000 horas?

d) Un ingeniero, modificando el proceso de fabricación del circuito, descubre que con sus cambios, la probabilidad de que un circuito dure al menos 1000 horas, es de 3/4, ¿qué decisión tomaría usted respecto al proceso existente?

e) Halle, si es posible, E(X).

10. La función de densidad de la variable aleatoria continua X es:

$$f(x) = \begin{cases} 4x(9-x^2) / 81 & 0 \leq x \leq 3 \\ 0 & \text{en otro caso} \end{cases}$$

Calcule e interprete:

a) El valor esperado.

b) La mediana.

c) El valor modal.

d) La varianza.

e) El cuartil superior.

f) El percentil 43.

11. La duración, en minutos, de una llamada telefónica de larga distancia se asimila en una variable aleatoria X con función de distribución:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - (2/3).e^{(-2/3)\,x} - (1/3).e^{(-1/3)\,x} & x > 0 \end{cases}$$

Calcule:

a) f(x)

b) la probabilidad de que una llamada dure menos de 2 minutos o más de 10 minutos.

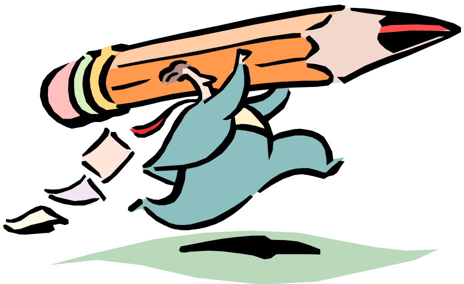
¡A repasar...!

Sabemos que ha encarado solo este tema y que puede tener algunas dudas.

Es preciso que lo domine antes de comenzar las aplicaciones prácticas y las autoevaluaciones.

Para autoevaluarse, responda las preguntas que están a continuación. Puede hacerlo con el material de estudio, pero asegurándose que "entiende" cada palabra, a tal punto que usted podría explicarle a un amigo, que no conoce el tema, de manera simple, los conceptos estudiados:

- ☑ **¿Qué es una variable aleatoria?**
- ☑ **¿Qué tipo de variables aleatorias conoce?**
- ☑ **¿Se trabajan y estudian algebraicamente de la misma manera los distintos tipos de variables?**
- ☑ **¿Qué funciones describen a una variable aleatoria y de qué manera lo hacen?**
- ☑ **¿Qué propiedades debe cumplir una función para ser función de densidad de una variable discreta?**
- ☑ **¿Qué propiedades debe cumplir una función para ser función de densidad de una variable continua?**
- ☑ **¿Qué propiedades debe cumplir una función para ser función de distribución de una variable discreta?**
- ☑ **¿Qué propiedades debe cumplir una función para ser función de distribución de una variable continua?**
- ☑ **¿Cómo se pasa de una función a otra?**
- ☑ **¿Qué representa cada una de las funciones en cada tipo de variable aleatoria?**
- ☑ **¿Es capaz de distinguir la forma de representar gráficamente las distintas funciones para los distintos tipos de variables?**
- ☑ **¿Es capaz de identificar correctamente, qué se representa en el eje de ordenadas, de abscisas y área bajo la curva en cada caso?**
- ☑ **¿Qué es la esperanza, valor esperado o media de una variable aleatoria (fórmula)? ¿Cómo se interpreta?**
- ☑ **¿Qué es la varianza de una variable aleatoria (fórmula)? ¿Cómo se interpreta?**
- ☑ **Enuncie las propiedades de esperanza y varianza.**
- ☑ **¿Es la esperanza un operador lineal?, ¿y la varianza?**
- ☑ **Enuncie el teorema de Chebyshev.**
- ☑ **¿Para qué sirve el teorema de Chebyshev?**



Por favor, no avance al siguiente tema si tiene dudas o no recuerda las nociones aquí volcadas. Pero si se siente listo para continuar, es hora de empezar a trabajar con las **autoevaluaciones** y las **aplicaciones prácticas**...

Ejercicios resueltos

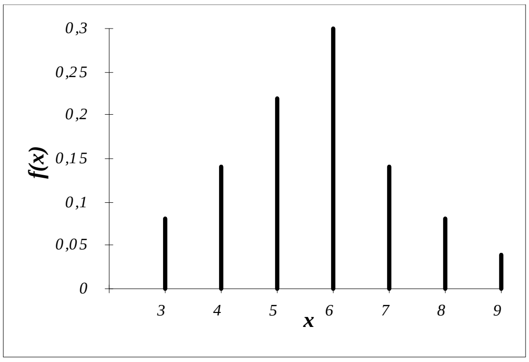
1. Clasifique las siguientes variables aleatorias como discretas o continuas:

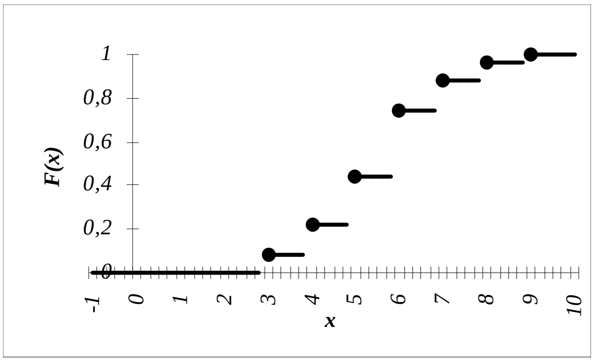
A: Producción diaria de motores en una empresa automotriz.	*Discreta*
B: Tiempo de espera hasta acceder a un servidor.	*Continua*
C: Venta, en dólares, de cada vendedor de un comercio.	*Continua*
D: Vida útil de un rodamiento.	*Continua*
E: Cantidad de unidades defectuosas en una línea de producción.	*Discreta*
F: Nivel de contaminación ambiental.	*Continua*
G: Cantidad de intentos hasta acertar en un blanco.	*Discreta*
H: Sueldo de los empleados de cierta industria.	*Continua*

2. El grado con el cual una compañía se involucra con el desarrollo de nuevos productos depende, en gran medida, de sus competidores. Un estudio fue realizado sobre las compañías que utilizan tecnología avanzada, para analizar el número de nuevos productos que se lanzan al mercado anualmente. Los cálculos, en función de años anteriores, están dados en la siguiente tabla:

X	3	4	5	6	7	8	9
f(x)	0,08	0,14	0,22	0,30	0,14	0,08	0,04
F(x)	*0,08*	*0,22*	*0,44*	*0,74*	*0,88*	*0,96*	*1,00*

a) Grafique la función de cuantía y la función de probabilidad acumulada.





b) Calcule el valor esperado, la varianza y la desviación estándar.

E(X) $= 3.0{,}08 + 4.0{,}14 + 5.0{,}22 + 6.0{,}30 + 7.0{,}14 + 8.0{,}08 + 9.0{,}04 = \mathbf{5{,}68}$

Var(X) $= E(X^2) - (E(X))^2 = 34{,}48 - (5{,}68)^2 = \mathbf{2{,}2176}$

$E(X^2) = 3^2.0{,}08 + 4^2.0{,}14 + 5^2.0{,}22 + 6^2.0{,}30 + 7^2.0{,}14 + 8^2.0{,}08 + 9^2.0{,}04 =$

$= \mathbf{34{,}48}$

D.E.(X) $= \sqrt{2{,}2176} = \mathbf{1{,}4892}$

c) Encuentre la proporción de compañías de tecnología avanzada para las cuales la variable aleatoria iguala o excede a $\mu + 2\sigma$.

*La variable aleatoria iguala o excede a $\mu + 2\sigma = 5{,}68 + 2 \,.\, 1{,}4892 = 8{,}6583$ en $x = 9$, luego, la probabilidad de que $X \geq 9$ es de 0,04, o sea, el **4%** de las compañías de tecnología avanzada lanza al mercado anualmente, una cantidad de 9 o más nuevos productos.*

3. Una caja contiene tres bolillas numeradas del 1 al 3, primero se extrae una bolilla y luego se lanza una moneda legal tantas veces como indique la bolilla. Calcule f(x), F(x) y grafique, siendo X: 'número de caras obtenidas'.

X: 'número de caras obtenidas al lanzar una moneda legal'

Bolilla	Moneda	X	Bolilla	Moneda	X	Bolilla	Moneda	X
①	C	1		CC	2		CCC	3
	X	0	②	CX	1		CCX	2
				XC	1		CXC	2
				XX	0	③	XCC	2
							CXX	1
							XCX	1
							XXC	1
							XXX	0

$P(X = 0) = P(0) = P(\,0\,/\,①)\,P(①) + P(\,0\,/\,②)\,P(②) + P(\,0\,/\,③)\,P(③) =$

$= 1/2 \;.\; 1/3 \;+\; 2/4 \;.\; 1/3 \;+\; 1/8 \;.\; 1/3 \;=\; 7/24$

$P(X = 1) = P(1) = P(\,1\,/\,①)\,P(①) + P(\,1\,/\,②)\,P(②) + P(\,1\,/\,③)\,P(③) =$

$= 1/2 \;.\; 1/3 \;+\; 1/2 \;.\; 1/3 \;+\; 3/8 \;.\; 1/3 \;=\; 11/24$

$P(X = 2) = P(2) = P(\,2\,/\,①)\,P(①) + P(\,2\,/\,②)\,P(②) + P(\,2\,/\,③)\,P(③) =$

$= 0 \;.\; 1/3 \;+\; 1/4 \;.\; 1/3 \;+\; 3/8 \;.\; 1/3 \;=\; 5/24$

$P(X = 3) = P(3) = P(\,3\,/\,①)\,P(①) + P(\,3\,/\,②)\,P(②) + P(\,3\,/\,③)\,P(③) =$

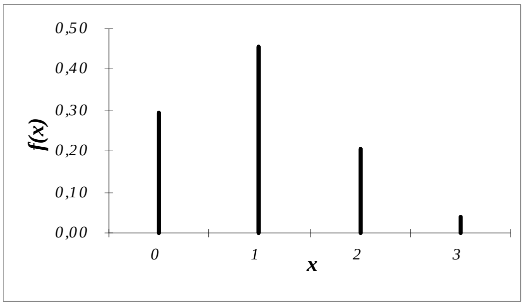
$= 0 \;.\; 1/3 \;+\; 0 \;.\; 1/3 \;+\; 1/8 \;.\; 1/3 \;=\; 1/24$

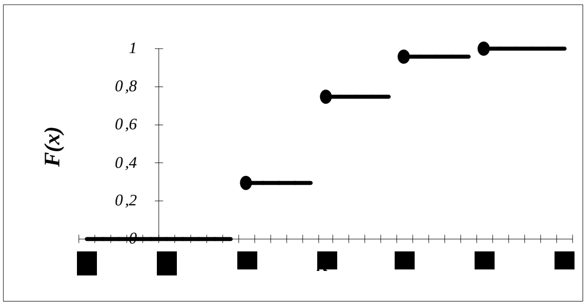
X	0	1	2	3
f(x)	7/24	11/24	5/24	1/24
F(x)	7/24	18/24	23/24	1

$E(X) = 0 \,.\, 7/24 + 1 \,.\, 11/24 + 2 \,.\, 5/24 + 3 \,.\, 1/24 \;=\; \mathbf{1}$

$Var(X) = E(X^2) - (E(X))^2 = 5/3 - 1^2 = \mathbf{2/3 = 0{,}67}$

$E(X^2) = 0^2 \,.\, 7/24 + 1^2 \,.\, 11/24 + 2^2 \,.\, 5/24 + 3^2 \,.\, 1/24 = 5/3$





4. Se lanza una moneda insesgada hasta que salga una cara o cinco cruces. Encuentre el número esperado de lanzamientos de la moneda.

X: 'número de lanzamientos'

ε	C	XC	XXC	XXXC	XXXXC	XXXXX
X	1	2	3	4	5	
f(x)	1/2	1/2.1/2=1/4	1/2.1/2.1/2=1/8	1/2.1/2.1/2.1/2=1/16	2.(1/2.1/2.1/2.1/2.1/2)=2/32	
F(x)	1/2	3/4	7/8	15/16	1	

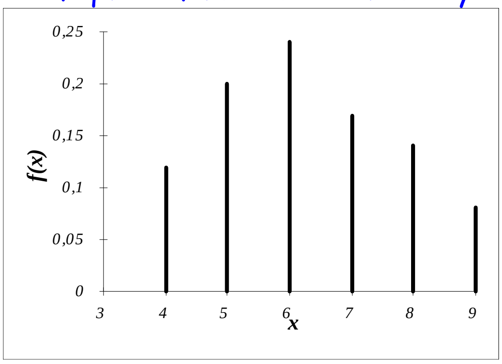
$E(X) = 1 \,.\, 1/2 + 2 \,.\, 1/4 + 3 \,.\, 1/8 + 4 \,.\, 1/16 + 5 \,.\, 2/32 = \mathbf{31/16 = 1{,}9375}$

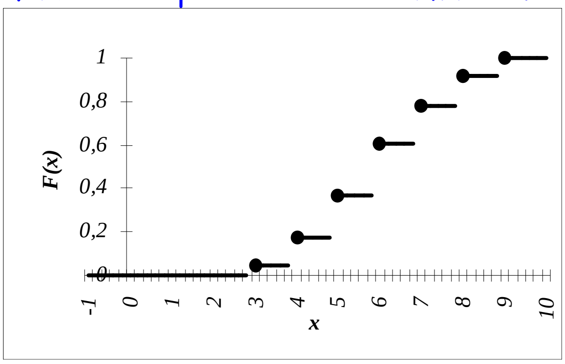
El número esperado de lanzamientos es, aproximadamente, de 2.

5. Se sabe por experiencia que la demanda diaria de un producto perecedero es como se muestra en la siguiente tabla:

X	3	4	5	6	7	8	9
f(x)	0,05	0,12	0,20	0,24	0,17	0,14	0,08
F(x)	0,05	0,17	0,37	0,61	0,78	0,92	1,00

a) Grafique la función de cuantía y la función de probabilidad acumulada.





b) Si cada artículo tiene un costo total de $45 con un precio de venta de $75, ¿cuál es la utilidad esperada?

$E(X) = 3.0{,}05 + 4.0{,}12 + 5.0{,}20 + 6.0{,}24 + 7.0{,}17 + 8.0{,}14 + 9.0{,}08 =$

= 6,1 productos

Se espera vender 6,1 productos por día con un costo total de 6,1 . \$45 = = \$274,5 que serán vendidos a \$75 cada uno, es decir, se recaudarán 6,1 . \$75 = \$457,5.

*Por lo tanto, la utilidad esperada será de \$457,5 - \$274,5 = **\$ 183**.*

6. Un jugador lanza dos monedas no cargadas. Gana \$2 si salen 2 caras y \$1 si sale 1 cara. Por otra parte, pierde \$3 si no sale ninguna cara. Determine el valor esperado del juego y si éste es favorable al jugador. Si el juego debe ser justo, ¿cuántos debería perder si no sale ninguna cara?

X: 'ganancia'

ε	CC	XC	CX	XX
X	**\$2**	**\$1**		**- \$3**
f(x)	1/2.1/2=1/4	2.1/2.1/2 =2/4		1/2.1/2=1/4

E(X)** = \$2 . 1/4 + \$1 . 2/4 + (-\$3) . 1/4 = **1 / 4 = \$ 0,25

Para que el juego sea justo la esperanza debe ser cero

ε	CC	XC	CX	XX
X	**\$2**	**\$1**		**\$?**
f(x)	1/2.1/2=1/4	2.1/2.1/2 =2/4		1/2.1/2=1/4

E(X) = \$2 . 1/4 + \$1 . 2/4 + y . 1/4 = \$ 0

*Entonces **y** = (- \$2 . 1/4 - \$1 . 2/4) / 1/4 = **- \$4***

Luego, para que el juego sea justo cuando no salgan caras debe perder \$4.

7. Se lanza una moneda al aire 3 veces. Si se obtienen al menos 2 caras se permite tirar un dado y recibir en dólares el número de puntos obtenido. ¿Qué es lo que puede esperar ganar en este juego en un intento?

X: 'ganancia'

ε		X
CCC	①	U\$S 1
	②	U\$S 2
	③	U\$S 3
	④	U\$S 4
	⑤	U\$S 5
	⑥	U\$S 6
CCX	①	U\$S 1
	②	U\$S 2
	③	U\$S 3
	④	U\$S 4
	⑤	U\$S 5
	⑥	U\$S 6
CXX	----	U\$S 0
XCX	----	U\$S 0

ε		X
CXC	①	U\$S 1
	②	U\$S 2
	③	U\$S 3
	④	U\$S 4
	⑤	U\$S 5
	⑥	U\$S 6
XCC	①	U\$S 1
	②	U\$S 2
	③	U\$S 3
	④	U\$S 4
	⑤	U\$S 5
	⑥	U\$S 6
XXC	----	U\$S 0
XXX	----	U\$S 0

Variable aleatoria

X	f(x)
U$S 0	4 . 1/8 = 4/8
U$S 1	4 . 1/6 . 1/8 = 4/48
U$S 2	4 . 1/6 . 1/8 = 4/48
U$S 3	4 . 1/6 . 1/8 = 4/48
U$S 4	4 . 1/6 . 1/8 = 4/48
U$S 5	4 . 1/6 . 1/8 = 4/48
U$S 6	4 . 1/6 . 1/8 = 4/48

E(X) *= U$S0 . 4/8 + U$S1 . 4/48 + U$S2 . 4/48 + U$S3 . 4/48 + U$S4 . 4/48 + U$S5 . 4/48 + U$S6 . 4/48 =* ***U$S 1,75***

Se puede esperar ganar U$S 1,75 en un intento.

8. Dada la variable aleatoria continua X con función de densidad $f(x) = k.x^2.(1-x)$ para $0 \leq x \leq 1$ y 0 para otro caso.

a) Determine la constante k.

$$\int_0^1 f(x)\,dx = \int_0^1 k.x^2.(1-x)\,dx = k.\int_0^1 (x^2-x^3)dx = k.\left[\int_0^1 x^2\,dx - \int_0^1 x^3\,dx\right] =$$

$$= k.\left[\left.\frac{x^3}{3}\right|_0^1 - \left.\frac{x^4}{4}\right|_0^1\right] = k.\left[\frac{1}{3}(1-0) - \frac{1}{4}(1-0)\right] = k.\frac{1}{12} = 1 \Rightarrow \mathbf{k = 12}$$

b) Halle F(x).

$$\mathbf{F(x)} = \int_0^x f(t)dt = \int_0^x k.t^2.(1-t)dt = \frac{1}{12}.\int_0^x (t^2-t^3)dt = \frac{1}{12}.\left[\int_0^x t^2dt - \int_0^x t^3dt\right] =$$

$$= \frac{1}{12}.\left[\left.\frac{t^3}{3}\right|_0^x - \left.\frac{t^4}{4}\right|_0^x\right] = \frac{1}{12}.\left[\frac{1}{3}(x^3-0) - \frac{1}{4}(x^4-0)\right] = \frac{1}{36}.x^3 - \frac{1}{48}.x^4$$

$$\Rightarrow \quad F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{1}{36}.x^3 - \dfrac{1}{48}.x^4 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

9. Sea la variable aleatoria continua X: "duración en horas de un circuito electrónico" con función de densidad:

$$f(x) = \begin{cases} 500 / x^2 & x \geq 500 \\ 0 & x < 500 \end{cases}$$

a) Pruebe que f(x) es función de densidad.

1º) $f(x) \geq 0$

2º) $\int_{500}^{\infty} f(x)\,dx = \int_{500}^{\infty} (500/x^2)\,dx = 500.\int_{500}^{\infty} x^{-2}\,dx = 500.\left(\left.\frac{x^{-1}}{-1}\right|_{500}^{\infty}\right) =$

$$= -500 \cdot \left(\frac{1}{\infty} - \frac{1}{500} \right) = \quad -500 \cdot \left(0 - \frac{1}{500} \right) = 1$$

*Luego, **f(x) es función de densidad.***

b) Halle F(x).

$$\int_{500}^{x} f(t)\,dt = \int_{500}^{x} (500/t^2)\,dt = 500 \int_{500}^{x} t^{-2}\,dt = 500. \left(\frac{t^{-1}}{-1} \Big|_{500}^{x} \right) =$$

$$= -500 \cdot \left(\frac{1}{x} - \frac{1}{500} \right) = -\frac{500}{x} + 1$$

$$\Rightarrow \qquad F(x) = \begin{cases} 0 & x < 500 \\ -\dfrac{500}{x} + 1 & x \geq 500 \end{cases}$$

c) ¿Cuál es la probabilidad de que un tubo dure entre 1500 y 2000 horas?

$$P(1500 < X < 2000) = F(2000) - F(1500) =$$

$$= -\frac{500}{2000} + 1 - \left(-\frac{500}{1500} + 1 \right) = -\frac{1}{4} + 1 + \frac{1}{3} - 1 = \frac{1}{12} = 0{,}08\widehat{3} \rightarrow 8{,}33\%$$

Hay un 8,33% de posibilidades de que el tubo dure entre 1500 y 2000 horas.

d) Un ingeniero, modificando el proceso de fabricación del circuito, descubre que con sus cambios, la probabilidad de que un circuito dure al menos 1000 horas, es de 3/4, ¿qué decisión tomaría usted respecto al proceso existente?

Según el proceso existente, la probabilidad de que el circuito dure al menos 1000 horas es

$$P(X > 1000) = 1 - F(1000) = 1 - \left(-\frac{500}{1000} + 1 \right) = 1 + \frac{1}{2} - 1 = 0{,}5 \rightarrow 50\%$$

Si modificando el proceso la probabilidad de que el circuito dure al menos 1000 horas es de 0,75, entonces conviene, en principio, modificar el proceso.

e) Halle, si es posible, E(x).

$$E(X) = \int_{500}^{\infty} x.f(x)\,dx = \int_{500}^{\infty} x.(500/x^2)\,dx = 500. \int_{500}^{\infty} x^{-1}\,dx =$$

$$= 500. \left(\ln|x| \,\Big|_{500}^{\infty} \right) = 500.(\ln\infty - \ln 500)$$

No se puede calcular la esperanza porque ln ∞ no se puede determinar.

10. La función de densidad de la variable aleatoria continua X es:

$$f(x) = \begin{cases} 4x(9-x^2)/81 & 0 \leq x \leq 3 \\ 0 & \text{en otro caso} \end{cases}$$

Calcule:

a) El valor esperado

$$\mathbf{E(X)} = \int_0^3 x\, f(x)\, dx = \int_0^3 x\,(4x(9-x^2)/81)\, dx = \int_0^3 (4x^2\,(9-x^2)/81)\, dx =$$

$$= (4/81)\int_0^3 x^2\,(9-x^2)\, dx = \mathbf{8/5 = 1,6}$$

El valor esperado es 1,6.

b) La mediana

La mediana es un valor a, tal que $P(X \le a) = \frac{1}{2}$

$$P(X \le a) = \int_0^a f(x)\, dx = \int_0^a 4x(9-x^2)/81\, dx = (4/81)\int_0^a x\,(9-x^2)\, dx =$$

$$= (4/81)\int_0^a (9x - x^3)\, dx = \frac{4}{81}\cdot\left(\frac{9a^2}{2} - \frac{a^4}{4}\right) = \frac{1}{2}$$

Luego:

$$\frac{4}{81}\cdot\left(\frac{9a^2}{2} - \frac{a^4}{4}\right) - \frac{1}{2} = 0 \Rightarrow \frac{2a^2}{9} - \frac{a^4}{81} - \frac{1}{2} = 0 \Rightarrow$$

$\Rightarrow$ *(multiplicando por -2/81) se llega a la expresión* $2a^4 - 36a^2 + 81 = 0$ *y podemos expresar la ecuación bicuadrática como una ecuación cuadrática de la forma:* $2(a^2)^2 - 36a^2 + 81 = 0$*. Luego, haciendo* $a^2 = x$*, queda la forma de la tradicional ecuación cuadrática:* $2x^2 - 36x + 81 = 0$

de donde $x = a^2 = \dfrac{36 \pm \sqrt{36^2 - 4.2.81}}{2.2} = \dfrac{36 \pm \sqrt{648}}{4} = 9 \pm \dfrac{9}{2}\sqrt{2}$

por lo tanto, el valor de la mediana, denominado en este caso con la letra a, será $a = \sqrt{9 \pm \dfrac{9}{2}\sqrt{2}}$ *y como la mediana debe estar entre 0 y 3, porque es en este intervalo donde está definida la variable, tomamos el valor* $\mathbf{a} = \sqrt{9 - \dfrac{9}{2}\sqrt{2}}$ *, de donde* $\mathbf{Me \approx 1,6236}$

c) El valor modal

El modo se obtiene haciendo máxima a f(x), es decir,

$$\frac{d}{dx}\left[\frac{4x\,(9 - x^2)}{81}\right] = \frac{36 - 12\,x^2}{81} = 0 \Rightarrow x = \sqrt{3} \approx \pm 1,732$$

la derivada segunda es $\dfrac{-24\,x}{81}$ *y para* $x \approx +1,732$*, la derivada segunda* $\dfrac{-24\,x}{81} < 0$*, por lo tanto este valor es un máximo, luego,* $\mathbf{Mo \approx +1,732}$

El modo es el valor 1,732, es decir, es el valor que se presenta con mayor frecuencia.

d) La varianza

$$E(X^2) = \int_0^3 x^2\, f(x)\, dx = \int_0^3 x^2\,(4x(9-x^2)/81)\, dx = \int_0^3 (4x^3\,(9-x^2)/81)\, dx =$$

$= (4/81) \int_0^3 x^3 (9-x^2)\, dx = 3$

Luego, $\mathbf{Var(X)} = E(X^2) - (E(X))^2 = 3 - (8/5)^2 = 3 - (64/25) = \mathbf{0{,}44}$

La varianza es de 0,44.

e) El cuartil superior

El cuartil 3 es el valor a, tal que $P(X \le a) = 0{,}75$

$P(X \le a) = \int_0^a f(x)\, dx = \int_0^a 4x(9-x^2)/81\, dx = (4/81)\int_0^a x(9-x^2)\, dx =$

$= (4/81)\int_0^a (9x - x^3)\, dx = \dfrac{4}{81}\cdot\left(\dfrac{9a^2}{2} - \dfrac{a^4}{4}\right) = \dfrac{3}{4}$

Luego

$\dfrac{4}{81}\cdot\left(\dfrac{9a^2}{2} - \dfrac{a^4}{4}\right) - \dfrac{3}{4} = 0 \Rightarrow \dfrac{2a^2}{9} - \dfrac{a^4}{81} - \dfrac{3}{4} = 0 \Rightarrow$

se llega a la expresión $-\dfrac{1}{81}a^4 + \dfrac{2}{9}a^2 - \dfrac{3}{4} = 0$

de donde $\mathbf{Q_3 \approx 2{,}1213}$

El cuartil superior es el valor 2,1213.

f) El percentil 43

El percentil 43 es el valor a, tal que $P(X \le a) = 0{,}43$

$P(X \le a) = \int_0^a f(x)\, dx = \int_0^a 4x(9-x^2)/81\, dx = (4/81)\int_0^a x(9-x^2)\, dx =$

$= (4/81)\int_0^a (9x - x^3)\, dx = \dfrac{4}{81}\cdot\left(\dfrac{9a^2}{2} - \dfrac{a^4}{4}\right) = \dfrac{43}{100}$

Luego

$\dfrac{4}{81}\cdot\left(\dfrac{9a^2}{2} - \dfrac{a^4}{4}\right) - \dfrac{43}{100} = 0 \Rightarrow \dfrac{2a^2}{9} - \dfrac{a^4}{81} - \dfrac{43}{100} = 0 \Rightarrow$

se llega a la expresión $-\dfrac{1}{81}a^4 + \dfrac{2}{9}a^2 - \dfrac{43}{100} = 0$

de donde $\mathbf{P_{43} \approx 1{,}4850}$

El percentil 43 toma el valor 1,4850.

11. La duración, en minutos, de una llamada telefónica de larga distancia se asimila en una variable aleatoria X con función de distribución:

$$F(x) = \begin{cases} 0 & x \le 0 \\ 1 - (2/3).e^{(-2/3)x} - (1/3).e^{(-1/3)x} & x > 0 \end{cases}$$

Calcule:

a) La función de densidad

$f(x) = F'(x) = 0 - (2/3)(-2/3).e^{(-2/3)\,x} - (1/3).(-1/3).e^{(-1/3)\,x} =$
$= (4/9).e^{(-2/3)\,x} + (1/9).e^{(-1/3)\,x}$, entonces

$$f(x) = \begin{cases} 0 & x \leq 0 \\ (4/9).e^{(-2/3)\,x} + (1/9).e^{(-1/3)\,x} & x > 0 \end{cases}$$

b) La probabilidad de que una llamada dure menos de 2 minutos o más de 10 minutos.

$\boldsymbol{P(X \leq 2) + P(X > 10)} = F(2) + (1 - F(10)) =$
$= 1 - (2/3).e^{(-2/3)\,2} - (1/3).e^{(-1/3)\,2} + 1 - 1 + (2/3).e^{(-2/3)\,10} + (1/3).e^{(-1/3)\,10} =$
$= \mathbf{0{,}6659} \rightarrow \mathbf{66{,}59\%}$

La probabilidad de que una llamada dure menos de 2 minutos o más de 10 minutos es de 0,6659.