Bloque I(primera semana): Experimentos aleatorios: espacio muestral, eventos. Definición axiomática de probabilidad. Propiedades derivadas. Probabilidad Condicional, regla del producto, independencia. Variables aleatorias discretas y continuas. Funciones de densidad, cuantía y distribución. Esperanza y Varianza. Distribuciones especiales: binomial, geométrica, Poisson, Hipergeométrica, normal, exponencial y uniforme.

El estudio de la probabilidad surgió hace tres siglos como una rama de la matemática que se originó en los juegos de azar.

***Definiciones*.**

I.- Llamamos ***experimento*** a cualquier proceso o acción que genera observaciones o resultados.
Por ejemplo: "tirar una moneda", "seleccionar un individuo y pesarlo", "observar el clima en un día determinado", "estudiar el aumento o disminución de una acción de un día al siguiente".

II.-El ***Espacio muestral*** asociado a un experimento, es el conjunto de todos los resultados posibles de dicho experimento.
Se denota generalmente con **Ω ; *E*** ó ***S***.

Ejemplos:

1.- Experimento: "arrojar una moneda al aire y registrar su resultado"
Espacio muestral: $\Omega = \{cara; ceca\}$; $\Omega = \{1 ; 0\}$; si aposté a que salga cara puedo pensar el espacio muestral como $\Omega = \{$ éxito ; fracaso$\}$.

2.- Experimento: "arrojar una moneda al aire dos veces y registrar el resultado"
Espacio muestral: $\Omega = \{ (cara, cara) ; (cara, ceca) ; (ceca, cara) ; (ceca, ceca) \}$ ó bien $\Omega = \{ (1,1) ; (1, 0) ; (0, 1) ; (0, 0) \}$

3.- Experimento: "arrojar una moneda al aire dos veces y registrar el número de caras"
Espacio muestral: $\Omega = \{0 ; 1 ; 2\}$

4.- Experimento: ensayo un procedimiento hasta que logro un primer éxito por
Espacio muestral: $\Omega = \{E; FE; FFE; FFFE; FFFFE; \ldots\}$
Por ejemplo: "arrojar una moneda al aire hasta que aparece cara por primera vez"
Espacio muestral: $\Omega = \{1 ; 01 ; 001 ; 0001 ; 00001 ; \ldots\} = \{ (x_1, x_2,\ldots, x_n): n\in \mathbf{N} ; x_i = 0 \ \forall\, i < n ; x_n = 1\} = \mathbf{N}$.

5.- Experimento: se enciende una lamparilla y se mide el tiempo transcurrido hasta que se quema.
Espacio muestral: $\Omega = \mathbf{R}^+ = \mathbf{R}_{>0} = (0, +\infty)$.

6.- Experimento: se arrojan dos dados simultáneamente, uno rojo y otro azul, yse observan los pares de resultados obtenidos.
Espacio muestral: $\Omega = \{ (i , j) : 1 \leq i \leq 6 ; 1 \leq j \leq 6\}$

Observar que el espacio muestral asociado a un experimento puede ser finito (como en los ejemplos 1, 2, 5, 6) , infinitos numerables (como en el ejemplo 3) o infinito no numerable (como en el ejemplo 4).

Sucesos o eventos:

Cuando se realiza un experimento, no sólo interesan los resultados individuales posibles, sino que pueden interesarnos familias o colecciones de ellos.

Llamamos ***evento o suceso*** a un subconjunto de Ω.
Cuando Ω es finito o infinito numerable todo subconjunto de Ω es un suceso.
Si Ω es infinito no numerable "casi todo" subconjunto de Ω es un evento.

Definimos como **un *evento simple ó elemental*** a aquél que consiste de un solo resultado y como un *evento compuesto* al que consiste de más de un evento elemental.

EJEMPLOS DE EVENTOS:

1.- Experimento: "arrojar una moneda al aire y registrar su resultado"
Suceso: A = "sale una cara". A = {cara} *evento elemental*

3.- Experimento: "arrojar una moneda al aire dos veces y registrar el número de caras"
Evento: A = "sale a lo sumo 1 cara" A = {0, 1} *evento compuesto*

4.- Experimento: "arrojar una moneda al aire hasta que aparece cara por primera vez"
Suceso: A = "el número de tiros es impar" A = { 1 ; 001; 00001; 0000001; ....} = { $2k + 1$ $k \in$ **N**$\cup$**{0}**} *evento compuesto*

5.- Experimento: se enciende una bombita de electricidad y se mide el tiempo transcurrido hasta que se quema .
Evento: A = "la lámpara dura más de 25 horas" A = (25, + ∞) *evento compuesto*

6.- Experimento: se arrojan dos dados simultáneamente, uno rojo y otro azul, y se observan los resultados obtenidos.
Suceso: A = "se registra el mismo número en los dos dados" A = {(1,1);(2, 2);(3, 3);(4, 4); (5, 5); (6, 6)} *evento compuesto*

evento compuesto

Sean ***A*** y ***B*** sucesos asociados a un espacio muestral Ω, entonces a partir de ellos puede definirse la ocurrencia de eventos combinados.

3.- ***A*** $\cup$ ***B*** es el evento que ocurren los eventos ***A*** ,***B*** o ambos.
4.- ***A*** $\cap$ ***B*** es el evento que ocurre si ocurren simultáneamente el suceso ***A*** y el suceso ***B***.
5.- A^c ó $\overline{A}$ es el suceso que ocurre cuando no se verifica el suceso ***A***, se denomina ***suceso opuesto*** o **complementario**.
6.- ***A*** - ***B*** es el evento que ocurre si se verifica el suceso ***A*** y no se verifica el suceso ***B***, puede expresarse como una interseccion: $A \cap \overline{B}$.
7.- Se dice que ***A conduce a B*** (***A*** $\subseteq$ ***B***),o bien **A** esta contenido en **B** si la realización de ***A*** implica la realización de ***B***.

8.- Los sucesos ***A*** y ***B*** son ***mutuamente excluyentes*** si no pueden ocurrir simultáneamente, es decir su intersección es el suceso imposible. Simbólicamente se expresa: $\boldsymbol{A} \cap \boldsymbol{B} = \varnothing$.

Axiomas de Probabilidad:

Dado un experimento y un espacio muestral Ω asociado al experimento el objetivo de la teoría de la probabilidad es asociar a cada evento ***A*** un número $P(\boldsymbol{A})$, llamado probabilidad de ***A***, que de una medida precisa de la chance o probabilidad de que ocurra el evento ***A***.

Axiomas de Probabilidad:

1.- $P(A) \geq 0 \quad \forall$ evento *A* del espacio muestral Ω.

2.- $P(\Omega) = 1$

3.- Si $A_1, A_2, \ldots, A_n$ es una familia de eventos mutuamente excluyentes

$$(A_i \cap A_j \neq \varnothing \ \forall \ i \neq j) \Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

Propiedades de la Probabilidad:

1.- $P(\boldsymbol{A}) = 1 - P(\boldsymbol{A}^C) \quad \forall$ evento ***A*** del espacio muestral Ω.
DEMOSTRACIÓN

$$A \cap A^C = \varnothing \text{ y } A \cup A^C = \Omega \Rightarrow P\left(A \cup A^C\right) = P(\Omega) = 1 \Rightarrow 1 = P\left(A \cup A^C\right) = P(A) + P\left(A^C\right) \Rightarrow P\left(A^C\right) = 1 - P(\boldsymbol{A})$$

2.- $P(\varnothing) = 0$
DEMOSTRACIÓN

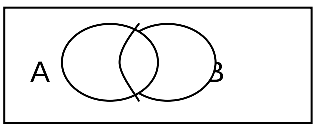
$$\varnothing \cap \varnothing^C = \varnothing \text{ y } \varnothing \cup \varnothing^C = \Omega \ \therefore \varnothing^C = \Omega - \varnothing = \Omega \Rightarrow 1 = P(\varnothing) + P\left(\varnothing^C\right) \text{ y}$$

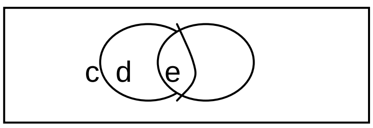
$$P\left(\varnothing^C\right) = P(\Omega) = 1 \Rightarrow P(\varnothing) = 1 - P\left(\varnothing^C\right) = 1 - 1 = 0 \Rightarrow P(\varnothing) = 0.$$

3.- Si $\boldsymbol{A} \subseteq \boldsymbol{B} \Rightarrow P(\boldsymbol{A}) \leq P(\boldsymbol{B})$
DEMOSTRACIÓN

$\boldsymbol{B} = A \cup \left(A^C \cap B\right)$ y $A \cap \left(A^C \cap B\right) = \varnothing \Rightarrow P(B) = P(A) + P\left(A^C \cap B\right)$. Como $P\left(A^C \cap B\right) \geq 0 \Rightarrow P(B) \geq P(A)$

4.- Si ***A*** y ***B*** son dos eventos cualesquiera de un espacio muestral Ω
$\Rightarrow P(\boldsymbol{A} \cup \boldsymbol{B}) = P(\boldsymbol{A}) + P(\boldsymbol{B}) - P(\boldsymbol{A} \cap \boldsymbol{B})$





DEMOSTRACIÓN

$\boldsymbol{A} \cup \boldsymbol{B} = A \cup \left(B \cap A^C\right)$ (mutuamente excluyentes) $\Rightarrow$ P($\boldsymbol{A} \cup \boldsymbol{B}$) = P(**A**)+P$\left(B \cap A^C\right)$

Además: $\boldsymbol{B} = \left(B \cap A\right) \cup \left(B \cap A^C\right)$ (excluyentes) $\Rightarrow$ P(**B**) = $P\left(B \cap A\right) + P\left(B \cap A^C\right)$

$\Rightarrow P\left(B \cap A^C\right)$ = P(**B**) - $P\left(B \cap A\right)$

De las igualdades: P($\boldsymbol{A} \cup \boldsymbol{B}$) = P(**A**)+P$\left(B \cap A^C\right)$ y $P\left(B \cap A^C\right)$ = P(**B**) - $P\left(B \cap A\right)$ se deduce, reemplazando, que: P($\boldsymbol{A} \cup \boldsymbol{B}$) = P(**A**) + P(**B**) - $P\left(B \cap A\right)$.

***Ejemplo* 1:**

En un barrio el 60% de las familias están suscriptas al periódico metropolitano, el 80% al periódico local y el 50% a ambos. Se selecciona una familia al azar:

a.- ¿Cuál es la probabilidad de que se encuentre suscripta a al menos uno de estos periódicos?

b.- ¿Cuál es la probabilidad de que se encuentre suscripta sólo a un periódico?

Sea ***A*** = “suscripta al periódico metropolitano” y ***B*** = “suscripta al periódico local”

Entonces: P(***A***) = 0.6 ; P(***B***) = 0.8 y P($\boldsymbol{A} \cap \boldsymbol{B}$) = 0.5

a.- debemos calcular P($\boldsymbol{A} \cup \boldsymbol{B}$) = P(***A***) + P(***B***) - P($\boldsymbol{A} \cap \boldsymbol{B}$) = 0.6 + 0.8 – 0.5 = 0.9

b.- debemos calcular P(***A***-***B***) + P(***B***- ***A***) = (P(***A***) - P($\boldsymbol{A} \cap \boldsymbol{B}$)) + (P(***B***) - P($\boldsymbol{A} \cap \boldsymbol{B}$)) =
= (0.6 – 0.5) + (0.8 – 0.5) = 0.4

Ejemplo 2:

Supongamos que arrojamos un dado equilibrado, calcular la probabilidad de obtener un resultado par.

Si el dado es equilibrado, podemos suponer que P(1) = P(2) = ...= P(6) = $\frac{1}{6}$

A = {2 ; 4 ; 6 } $\Rightarrow$ P(*A*) = $3\frac{1}{6} = \frac{1}{2}$

Ejemplo 3:

Supongamos que arrojamos un dado que no es equilibrado y que las caras pares tienen el doble de probabilidad que las impares, calcular la probabilidad de que el resultado sea par.

$p(1) = p(3) = p(5) = a$; P(2) = P(4) = P(6) = 2a como $\sum_{i=1}^{6} p(i) = 1 \Rightarrow$ 3a + 3.2a = 1

$\Rightarrow$ 9a = 1 $\Rightarrow a = \frac{1}{9}$. Por lo tanto $p(2) = p(4) = p(6) = 2a = \frac{2}{9} \Rightarrow$ P(A) = 3. $\frac{2}{9} = \frac{2}{3}$

Equiprobabilidad:

Sea un experimento aleatorio con *n* resultados posibles, es decir: #(Ω) = *n*. Diremos que el espacio muestral es de equiprobabilidad si: $p_i = p = P(e_i) \quad \forall\, e_i \in \Omega$.

De esta definición se deduce que $p = 1/\#(\Omega)$ pues:

$1 = P(\Omega) = \sum_{i=1}^{n} P(e_i) = \sum_{i=1}^{n} p = n.p \quad \Rightarrow \quad p = \frac{1}{n} =$

Si A es un evento cualquiera $\Rightarrow$ P(***A***) = #(A)/ #(Ω)

Es decir que en los espacios de equiprobabilidad, la probabilidad de un evento es igual a la proporción de puntos del espacio muestral que tal evento contiene.

Ejemplo 4:

De una urna que contiene 3 bolitas Rojas y 4 bolitas Blancas, se extrae una bolita al azar. ¿Cuál es la probabilidad de que salga una Roja?

$\Omega = \{ B_1 ; B_2 ; B_3 ; B_4 ; R_1 ; R_2 ; R_3 \}$ $\quad A = (\{R_1 ; R_2 ; R_3 \}$

$\#(\Omega) = 7 \;;\; \#(A) = 3 \;\Rightarrow$

$P(B_i) = P(R_j) = \frac{1}{7} \quad \forall\, 1 \le i \le 4 \;;\; 1 \le j \le 3 \;\Rightarrow$

$P(R) = P(\{R_1 ; R_2 ; R_3 \}) = \frac{\# A}{\# \Omega} = \frac{3}{7}$

Probabilidad condicional:

Consideremos una caja que contiene 4 bolitas rojas de las cuales 2 son lisas y 2 rayadas, y 5 bolitas azules de las cuales 4 son lisas y 1 es rayada. Si se extrae una bolita al azar, ¿cuál es la probabilidad de que sea rayada?

Llamemos: ***A*** = "la bolita es rayada"

Considerando que es un espacio de equiprobabilidad: $P(A) = \frac{\# A}{\# \Omega} = \frac{3}{9} = \frac{1}{3}$

Imaginemos ahora, que hemos extraído una bolita de la caja sin mirarla y la persona que está a nuestro lado nos dice que es de color rojo. ¿Cuál es la probabilidad de que la bolita sea rayada?

Llamemos: ***A*** = "la bolita es rayada" y ***B*** = "la bolita es roja"

Como sabemos **B** es el suceso cierto (pues la bolita extraída es roja), podemos pensar que en realidad sólo extraje una bolita de una caja que contiene 4 bolitas rojas de las cuales 2 son lisas y dos son rayadas. Por lo tanto: P(***A***) = 0.5 (distinta de $\frac{1}{3}$).

Observación:

AL OCURRIR EL EVENTO ***B*** EL ESPACIO MUESTRAL SE REDUCE.

Sea P(***A***) la probabilidad del evento ***A***. Queremos ver cómo afecta a la probabilidad de ***A*** el hecho de conocer información sobre otro evento ***B***, queremos hallar esta probabilidad en función de P(***A***) y P(***B***).

Notación: $P\left(A/B\right)$ Se lee: "probabilidad condicional de ***A*** dado que ***B*** ocurrió" (Si ***B*** ocurrió entonces ***B*** no puede ser el suceso imposible y P(***B***) > 0).

Definición:

Sean ***A*** y ***B*** eventos tales que P(***B***) > 0 entonces definimos:

$$P\left(A/B\right) = \frac{P(A \cap B)}{P(B)}.$$

Ejemplo 5:

En el ejemplo de las bolitas rayadas y lisas:

P(***A***) = $\frac{1}{3}$; P(***B***) = $\frac{4}{9}$ y $P(A \cap B) = \frac{\# A \cap B}{\# \Omega} = \frac{2}{9}$ $\Rightarrow$ $P\left(A/B\right) = \frac{2/9}{4/9} = \frac{1}{2}$

Ejemplo 6:

En una población, la distribución de los portadores de HIV y su pertenencia a cierto grupo de riesgo que llamaremos I se muestra en la siguiente tabla:

		Portadores de HIV		
		SI	NO	
Pertenecen al Grupo de Riesgo I	SI	0.003	0.017	0.020
	NO	0.003	0.977	0.980
		0.006	0.994	1.000

Sean los sucesos: ***A*** = "es portador de HIV" y ***B*** = "pertenecen al Grupo de Riesgo I"

1.- En esta población, ¿cuál es la probabilidad de ser portador de HIV?
P(***A***) = 0.06

2.- ¿Cuál es la probabilidad de ser portador de HIV cuando una persona pertenece al Grupo de Riesgo I?

$$P\left(A/B\right) = \frac{P(A \cap B)}{P(B)} = \frac{0.003}{0.020} = 0.15$$

3.- ¿Cuál es la probabilidad de ser portador de HIV cuando una persona no pertenece al Grupo de Riesgo I?

$$P\left(A/\bar{B}\right) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{P(A \cap \bar{B})}{1 - P(B)} = \frac{0.003}{0.980} = 0.00306$$

(sólo 3 de cada mil personas de la población que no están en el Grupo de Riesgo I son portadores de HIV)

Regla del producto:

$$P(A \cap B) = P\left(A/B\right) . P(B)$$

Si P(***A***) > 0 $\Rightarrow$ $P(A \cap B) = P\left(B/A\right) . P(A)$

Ejemplo 7:

En la caja que tiene 4 bolitas Blancas y 3 bolitas Rojas, queremos calcular la probabilidad de extraer 1 Roja y 1 Blanca en ese orden, si la extracción se realiza sin reposición.

Sea ***A*** = "la primer bolita es Roja" y ***B*** = "la segunda bolita es Blanca"

Interesa calcular: $P(A \cap B) = P(A) \,.\, P\left(B/A\right) = \frac{3}{7} . \frac{4}{6} = \frac{12}{42} = \frac{2}{7}$

La regla del producto es muy útil cuando el experimento consiste en varias etapas sucesivas. Así si se desea calcular:

$$P(A_1 \cap A_2 \cap A_3) = P((A_1 \cap A_2) \cap A_3) = P(A_1 \cap A_2) \,.\, P\left(A_3 / A_1 \cap A_2\right)$$

$$= P(A_1) \,.\, P\left(A_2 / A_1\right) . \; P\left(A_3 / A_1 \cap A_2\right).$$

Esta regla se puede generalizar a n etapas, y permite calcular la probabilidad de la intersección de n eventos.

Propiedades de la probabilidad condicional:

Diagrama de árbol:

Ejemplo 8:

Para viajar de Buenos Aires a Córdoba hay tres compañías que operan en el horario deseado. La experiencia indica que la compañía A_1 sale con retraso con probabilidad 0.4, la compañía A_2 sale con retraso con probabilidad 0.5 y la A_3 sale con retraso con probabilidad 0.7. Si una persona selecciona una compañía al azar:

a.- ¿cuál es la probabilidad de que elija la compañía A_1 y salga con retraso?

Sea **R** = "sale con retraso"

Los datos indican: $P\left(R/A_1\right) = 0.4$; $P\left(R/A_2\right) = 0.5$; $P\left(R/A_3\right) = 0.7$ y $P(A_i) = \frac{1}{3}$

$\forall \; 1 \leq i \leq 3$ (elección al azar) $\Rightarrow$ $P(R \cap A_1) = P\left(R/A_1\right) . P(A_1) = 0.4 \,.\, \frac{1}{3} = \frac{2}{15}$

b.- ¿cuál es la probabilidad de que la persona salga con retraso?

$$P(R) = P\big((R \cap A_1) \cup (R \cap A_2) \cup (R \cap A_3)\big) =$$

$$= P(R \cap A_1) + P(R \cap A_2) + P(R \cap A_3) =$$

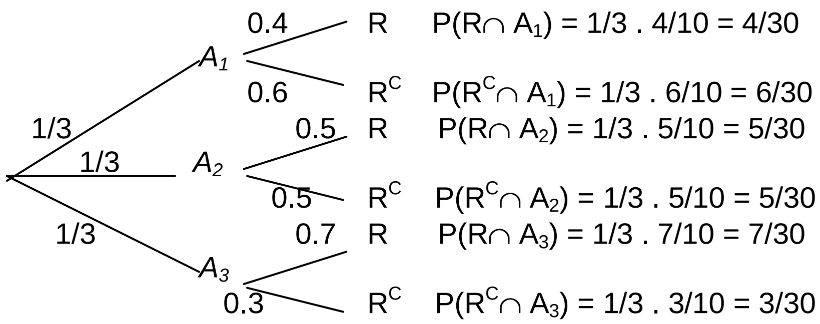
$$= P\left(R/A_1\right) . P(A_1) + P\left(R/A_2\right) . P(A_2) + P\left(R/A_3\right) . P(A_3) =$$

$$= 0.4 \,.\, \frac{1}{3} + 0.5 \, \frac{1}{3} + 0.7 \, \frac{1}{3} = 1.6 \, \frac{1}{3} = \frac{16}{30} = \frac{8}{15}$$

c.- dado que salgo con retraso, ¿cuál es la probabilidad de haber elegido la compañía A_1?

$$P\left(A_1/R\right) = \frac{P(A_1 \cap R)}{P(R)} = \frac{2/15}{8/15} = \frac{1}{4}$$

Para responder a estas preguntas se podría haber construido un diagrama de árbol de la forma siguiente:



Notar que $P(A_1) = \frac{1}{3}$ y $P\left(A_1/R\right) = \frac{1}{4}$. Es decir que cuando se sabe que la persona se retrasó en salir la probabilidad de haber elegido la compañía A_1 disminuye, esto se debe a que esta compañía es la más puntual de las tres.

Concepto de Partición:

Una colección de eventos $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, ..., $\boldsymbol{A_k}$ constituyen una partición del espacio Ω si:

i) $A_i \cap A_j = \varnothing \quad \forall\ i \neq j$ *ii*) $P(A_i) > 0\ \forall\ I$ *iii*) $\bigcup_{i=1}^{k} A_i = \Omega$

Teorema de la probabilidad total:

Sean $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, ..., $\boldsymbol{A_k}$ una partición del espacio Ω y sea $\boldsymbol{B}$ cualquier evento, entonces:

$$P(B) = \sum_{i=1}^{k} P\left(B/A_i\right).P(A_i)$$

Demostración:

$$\boldsymbol{B} = \boldsymbol{B} \cap \Omega = B \cap \left(\bigcup_{i=1}^{k} A_i\right) = \bigcup_{i=1}^{k} (B \cap A_i)\ \text{.Como}\ (B \cap A_i) \cap (B \cap A_j) = (B \cap A_i \cap A_j)$$

$$= B \cap \varnothing = \varnothing \Rightarrow P\left(\bigcup_{i=1}^{k} (B \cap A_i)\right) = \sum_{i=1}^{k} P(B \cap A_i) = \sum_{i=1}^{k} P\left(B/A_i\right).P(A_i)$$

Teorema de Bayes:

Sean $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, ..., $\boldsymbol{A_k}$ una partición del espacio Ω y sea $\boldsymbol{B}$ tal que P($\boldsymbol{B}$) > 0, entonces:

$$P\left(A_i/B\right) = \frac{P\left(B/A_i\right).P(A_i)}{\sum_{j=1}^{k} P\left(B/A_j\right).P(A_j)} \qquad \forall\ 1 \leq i \leq k$$

Demostración:

$$P\left(A_i/B\right) = \frac{P(A_i \cap B)}{P(B)} = \frac{P\left(B/A_i\right).P(A_i)}{P(B)} = \frac{P\left(B/A_i\right).P(A_i)}{\sum_{j=1}^{k} P\left(B/A_j\right).P(A_j)}$$ (usando el teorema de la probabilidad total)

Ejemplo 9:

Se sabe que uno de cada mil individuos contraen la enfermedad XX .
Para detectar esta enfermedad se usa un test que da resultado positivo en el 99% de los casos de personas enfermas, en tanto que da positivo sólo en el 2% de los casos de las personas sanas.

a.- Si tomamos un individuo al azar, le realizamos el test y da positivo, ¿cuál es la probabilidad de que haya contraído la enfermedad XX?
Sean: ***A*** = “el sujeto tiene la enfermedad” y **+** = “el resultado del test es positivo”

Datos: $P(A) = 0.001$; $P\left(+/A\right) = 0.99$; $P\left(+/\bar{A}\right) = 0.02$

Entonces: $P(\bar{A}) = 0.999$; $P\left(-/A\right) = 0.01$; $P\left(-/\bar{A}\right) = 0.98$

Se pide: $P\left(A/+\right)$ (probabilidad a posteriori de estar enfermo)

$$P\left(A/+\right) = \frac{P\left(+/A\right)P(A)}{P\left(+/A\right)P(A) + P\left(+/\bar{A}\right)P(\bar{A})} = \frac{0.99x0.001}{0.99x0.001 + 0.02x0.999} = 0.0472$$

Observar que la probabilidad a posteriori es aún muy baja, porque la probabilidad de contraer la enfermedad es muy pequeña.

b.- Si tomamos un individuo al azar, le realizamos el test y da negativo, ¿cuál es la probabilidad de que sea una persona sana?
Sean: **A** = “el sujeto tiene la enfermedad” y **-** = “el resultado del test es negativo”
Se pide:

$$P\left(\bar{A}/-\right) = \frac{P\left(-/\bar{A}\right)P(\bar{A})}{P\left(-/\bar{A}\right)P(\bar{A}) + P\left(-/A\right)P(A)} = \frac{0.98x0.999}{0.98x0.999 + 0.01x0.001} = 0.9999897$$

Independencia:

Hay casos en que $P\left(A/B\right) \neq P(A)$ como en el caso del ejemplo de HIV, es decir, que la información relacionada con la ocurrencia del evento ***B*** (pertenecer al grupo de riesgo I) modifica la probabilidad original de ser portador de HIV.

Sin embargo, hay situaciones en las que el tener información sobre la ocurrencia de otro evento no modifica $P(A)$.

Ejemplo 10:

Sea una caja que contiene 3 bolitas Rojas y 6 bolitas Azules. De las rojas 2 son lisas y 1 rayada, de las azules 4 son lisas y 2 rayadas. Se saca una bolita al azar y es Roja, ¿cuál es la probabilidad de que sea rayada?

Sean: ***A*** = “la bolita es rayada” y ***B*** = “la bolita es Roja”

Se sabe: $P(A) = \frac{1}{3}$; $P(B) = \frac{1}{3}$; $P(A \cap B) = \frac{1}{9}$

Se pide: $P\left(A/B\right) = \frac{P(A \cap B)}{P(B)} = \frac{1/9}{3/9} = \frac{1}{3} = P(A)$

La probabilidad de que la bolita sea rayada cuando sabemos que la bolita es roja es la misma que la probabilidad de ser rayada. Es decir, el hecho de conocer ***B*** no me da más información sobre la chance que tiene de ocurrir ***A***.

Decimos que dos eventos ***A*** y ***B*** son independientes si la información acerca de uno de ellos no modifica la probabilidad de que el otro ocurra.

Dos eventos ***A*** y ***B*** son independientes si: $P(A \cap B) = P(A).P(B)$

Ejemplo 11:

De un mazo de 40 cartas españolas, se extrae una carta al azar. Sean los sucesos: ***A*** = “la carta es de copa o espada” ; ***B*** = “la carta no es de copa” ; ***C*** = “la carta es de copa u oro”

a.- ¿Son ***A*** y ***B*** independientes?

Se sabe: $P(A) = \frac{1}{2}$; $P(B) = \frac{3}{4}$; $P(C) = \frac{1}{2}$

Las probabilidades correspondientes a estos eventos son:

	B	B^C	
A	1/4	1/4	1/2
A^C	2/4	0	1/2
	3/4	1/4	1

Si sabemos que ocurrió B^C (sale copa), entonces sabemos que ocurrió ***A*** pues $B^C \subseteq A$. No son sucesos independientes.

b.- ¿***A*** y ***C*** son eventos independientes?

Las probabilidades correspondientes a estos eventos son:

	C	C^C	
A	1/4	1/4	1/2
A^C	1/4	1/4	1/2
	1/2	1/2	1

A y C son independientes pues: $P(A \cap C) = \frac{1}{4} = \frac{1}{2}.\frac{1}{2} = P(A).P(C)$

Observar que: $P\left(A/C\right) = P(A) = \frac{1}{2}$ y $P\left(C/A\right) = P(A) = \frac{1}{2}$

Ejemplo 12:

De una caja que contiene 1 bolita Roja y 9 Blancas se extraen dos bolitas sucesivamente. Sean: R_1 = "la primer bolita es Roja" y R_2 = "la segunda bolita es Roja"

a.- Si las extracciones se realizan con reposición, ¿son R_1 y R_2 independientes?

$P(R_1) = \frac{1}{10}$; $P\left(R_2/R_1\right) = \frac{1}{10} \Rightarrow P\left(R_2/R_1\right) = P(R_2)$ o sea son sucesos independientes.

b.- Si las extracciones se realizan sin reposición, ¿son R_1 y R_2 independientes?

$P(R_1) = \frac{1}{10}$; $P\left(R_2/R_1\right) = 0 \Rightarrow P\left(R_2/R_1\right) \neq P(R_2)$ (no son eventos independientes)

Variables aleatorias discretas

El concepto de variable aleatoria permite pasar de los resultados de un experimento a valores numéricos asociados con dichos resultados.

Ejemplo 13:

Se arrojan tres monedas y se observa la secuencia de resultados.

Ω ={(c, c, c); (c, c, ce); (c ,ce ,c); (c, ce, ce); (ce, c, c); (ce, ce, c); (ce, c, ce); (ce, ce, ce)}

Podemos escribir un poco más sinteticamente :

$\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8\}$

Distintas variables aleatorias pueden asociarse a este experimento aleatorio:

X = "número de caras en los tres tiros"
Y = "número de caras en los dos primeros tiros"
Z = "número de caras menos número de cecas"

Definición formal:

Sea Ω un espacio muestral asociado a un experimento E. Una variable aleatoria es una función ***X*** que asocia a cada elemento *w* del espacio muestral Ω un número real que es ***X***(*w*).

Simbólicamente:

$$X: \Omega \rightarrow \mathbf{R} \quad \text{tal que} \quad x_i = X(w_i) \qquad \forall\, x_i \in \Omega$$

En el ejemplo 13:

$X((c, c, c)) = X(w_1) = 3$ (pues ***X*** cuenta el número de caras)
$Y((c, c, c)) = Y(w_1) = 2$ (pues ***Y*** cuenta el número de caras en los dos primeros tiros)
$Z((c, c, c)) = Z(w_1) = 3$ (pues ***Z*** cuenta el número de caras menos el número de cecas)

Para w_6 = (ce, ce, c): $\boldsymbol{X}(w_6) = 1$; $\boldsymbol{Y}(w_6) = 0$; $\boldsymbol{Z}(w_6) = -1$

Ejemplo 14:

a) Se arroja una moneda. Ω = {c, ce}. Si definimos $\boldsymbol{X}$(c)= 1; **X**(ce) = 0. El número 1 indica que salió cara y el 0 que salió ceca.

b) Si se tira un dado:
Ω = {1, 2, 3, 4, 5, 6,}. Podemos definir:

$\boldsymbol{X}(i) = 1$ si i es par , $\boldsymbol{X}(i) = 0$ si i es impar para todo $1 \le i \le 6$.

Esto puede expresarse designando acierto al evento si sale par

$\boldsymbol{Y}(i) = i$ (observo el número que sale)

Todas las variables de los ejemplos presentados corresponden a casos de variables discretas.
En los ejemplos anteriores, la variable aleatoria ***X*** toma sólo los valores 0 y 1 (arbitrariamente elegidos), en tanto que la variable ***Y*** toma seis valores distintos.

Una variable aleatoria es discreta si puede tomar un número finito o infinito numerable de valores.

En los ejemplos anteriores la variable es discreta pues toma un número finito de valores, veremos luego ejemplos de variables aleatorias discretas que toman un número infinito de valores.

Indicaremos con R_X **rango de la variable *X*** al conjunto de todos los valores posibles de ***X***.

Ejemplo 15:

1.- Se arroja un dado dos veces. Sean las variables aleatorias:

X = "suma los números obtenidos" R_X = {2, 3, 4, ..., 12}
Y = "toma el máximo entre los resultados de los dos dados" R_Y = {1, 2, 3,..., 6}

$\boldsymbol{Z} = \begin{cases} 1 & \textit{si la suma es } 6 \\ 0 & \textit{en caso contrario} \end{cases}$ $\qquad R_Z$ = {0, 1}

Z se denomina *variable de Bernoulli* pues sólo toma dos valores: 0 y 1.

2.- Se arroja un dado el número de veces necesaria hasta que se obtiene el primer as. Sea
Sea ***X*** = "número de tiros necesarios" $\qquad R_X = \boldsymbol{N}$

Definición:

La definición de probabilidad asociada a una variable discreta se define como:

$$p_X(x) = P(\boldsymbol{X} = x) = P(\{ w \in \Omega : \boldsymbol{X}(w) = x\}$$

De esta forma se verifica:

1) $p_X(x) \geq 0$ 2) $\sum_{x \in R_X} p_X(x) = 1$

La *función de probabilidad*, también llamada ***distribución de probabilidad*, *función de probabilidad puntual* o *función de probabilidad de masa*** de una variable aleatoria ***X*** nos caracteriza la variable aleatoria a partir de cómo se distribuye la probabilidad total 1 entre los distintos valores de la variable y se determina a partir de la probabilidad de los sucesos de Ω asociados a cada valor *x* de **X**.

Ejemplo 16:

E = "se lanza una moneda equilibrada dos veces", entonces: Ω = {(c, c); (c, ce); (ce, c); (ce, ce)} y $R_X = \{0, 1, 2\}$; sea ***X*** = "número de caras obtenidas"

$$p_X(0) = P(X = 0) = P(\{ w \in \Omega: \boldsymbol{X}(w) = 0\} = P[\{(ce, ce)\}] = \frac{1}{4}$$

$$p_X(1) = P(\boldsymbol{X} = 1) = P(\{ w \in \Omega: \boldsymbol{X}(w) = 1\} = P[\{(c, ce); (ce, c)\}] = \frac{1}{2}$$

$$p_X(2) = P(\boldsymbol{X} = 2) = P(\{ w \in \Omega: \boldsymbol{X}(w) = 2\} = P[\{(c, c)\}] = \frac{1}{4}$$

Si $x \neq 0, 1, 2 \Rightarrow p_X(x) = 0$. Por lo tanto: $\sum_{x \in R_X} p_X(x) = 1$

Podemos representar la función de probabilidad puntual en forma tabular, mediante un diagrama de líneas o con un histograma de probabilidad. Así en este caso:

x	0	1	2
P(**X** = *x*) = p(*x*)	0.25	0.5	0.25

Forma tabular

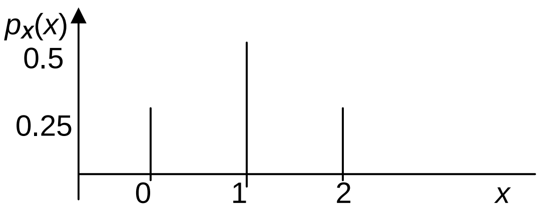
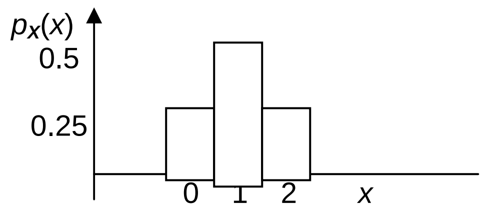


diagrama de líneas



histograma de probabilidad

Todos los rectángulos con base de igual long.

Función de distribución de probabilidades acumuladas:

En ocasiones dado un cierto *x* interesa conocer no sólo la probabilidad $p_X(x)$ sino la probabilidad asociada a valores de la variable menores o iguales que *x*, es decir: $p_X(\boldsymbol{X} \leq x)$. Esto da origen a nueva función que notaremos F(*x*) definida $\forall\, x \in$ **R.**

La función de ***distribución de probabilidad acumulada*** de una variable aleatoria discreta ***X*** con función de probabilidad puntual $p_X(x)$ se define para todo $x \in$ **R** como:

$$F_X(x) = p_X(\boldsymbol{X} \leq x) = \sum_{\substack{y \leq x \\ y \in R_X}} p_X(y)$$

O sea que la distribución de probabilidad acumulada nos da la probabilidad de que el valor observado sea a lo sumo x.

Ejemplo 17:

En el ejemplo de la moneda que se tira dos veces donde la probabilidad puntual de la variable $\boldsymbol{X}$ que cuenta el número de caras es:

x	0	1	2
$P(\boldsymbol{X} = x) = p(x)$	0.25	0.5	0.25

La función de distribución de probabilidad acumulada es:

Si $x < 0 \Rightarrow F(x) = F_X(x) = p_X(\boldsymbol{X} \leq x) = 0$

Si $x = 0 \Rightarrow F(0) = F_X(0) = p_X(\boldsymbol{X} \leq 0) = P(\boldsymbol{X} = 0) = 0.25$

Si $0 < x < 1 \Rightarrow F(x) = F_X(x) = p_X(\boldsymbol{X} \leq x) = P(\boldsymbol{X} = 0) = 0.25$

Si $x = 1 \Rightarrow F(1) = F_X(1) = p_X(\boldsymbol{X} \leq 1) = P(\boldsymbol{X} = 0) + P(\boldsymbol{X} = 1) = 0.75$

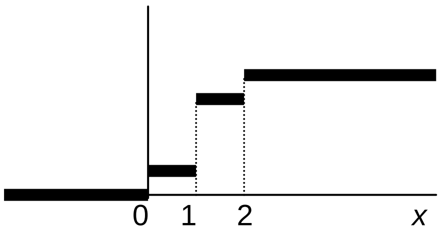
Si $1 < x < 2 \Rightarrow F(x) = F_X(x) = p_X(\boldsymbol{X} \leq x) = P(\boldsymbol{X} = 0) + P(\boldsymbol{X} = 1) = 0.75$

Si $x = 2 \Rightarrow F(2) = F_X(2) = p_X(\boldsymbol{X} \leq 2) = P(\boldsymbol{X} = 0) + P(\boldsymbol{X} = 1) + P(\boldsymbol{X} = 2) = 1$

Si $2 < x \Rightarrow F(x) = F_X(x) = p_X(\boldsymbol{X} \leq x) = P(\boldsymbol{X} = 0) + P(\boldsymbol{X} = 1) + P(\boldsymbol{X} = 2) = 1$

Resumiendo:

$$F_X(x) = \begin{cases} 0 & si\ x < 0 \\ 0.25 & si\ 0 \leq x < 1 \\ 0.75 & si\ 1 \leq x < 2 \\ 1 & si\ 2 \leq x \end{cases}$$



Observar que la función de distribución de probabilidad acumulada es una función que tiene las siguientes propiedades:

- es no decreciente
- toma valores entre 0 y 1
- es una función escalera, cuyo escalón final está a la altura 1 del eje vertical.

Proposición:

Sean a y b tales que $a \leq b$: $\quad P(a \leq x \leq b) = F(b) - F(a^-)$

(donde a^- representa el mayor valor de $\boldsymbol{X}$ estrictamente menor que a).

En particular: $P(X=a)=F(a)-F(a^-)$

Ejemplo:

En el caso de la moneda: $F_X(x) = \begin{cases} 0 & si\ x < 0 \\ 0.25 & si\ 0 \le x < 1 \\ 0.75 & si\ 1 \le x < 2 \\ 1 & si\ 2 \le x \end{cases}$

$P(1 \le x \le 2) = F(2) - F(1^-) = F(2) - F(0) = 1 - 0.25 = 0.75 = p(1) + p(2)$

$P(\boldsymbol{X} = 1) = F(1) - F(1^-) = F(1) - F(0) = 0.75 - 0.25 = 0.50 = p(1)$

Por lo tanto es posible reconstruir la probabilidad puntual, conociendo la función de distribución.

Parámetro de una función de probabilidad:

En un kiosco se observa que el 20% de los clientes pagan con billetes. Sea la variable:
$\boldsymbol{X}$ = “considera como éxito que el cliente pague con billetes”
$R_X = \{1, 0\}$ 1 = “cliente paga con billetes”; 0 = “cliente no paga con billetes”
$p_X(1) = 0.20$; $p_X(0) = 0.80$; $p_X(x) = 0 \ \forall x \neq 0$.

Si el comerciante no tuviese información sobre la probabilidad de que el cliente pague con billetes, podríamos escribir la función de probabilidad puntual como:

$p_X(1) = \alpha$; $p_X(0) = 1 - \alpha$; $p_X(x) = 0 \ \forall x \neq 0$ con $\alpha \in [0, 1]$

Suele notarse: $p_X(x, \alpha)$ para explicitar que esta función de probabilidad depende del parámetro α.

Definición:

Supongamos que $p_X(x)$ depende de una cantidad que puede tomar cualquier valor dentro de cierto rango. A dicha cantidad se la denomina ***parámetro de la distribución***.

La colección de todas las funciones de probabilidad para distintos valores del parámetro se denomina ***familia de distribuciones de probabilidad***.

Por ejemplo: la familia de distribuciones Bernoulli de parámetro α, con $0 < \alpha < 1$ es:

$$P_X(x, \alpha) = \begin{cases} \alpha & si\ x = 1 \\ 1 - \alpha & si\ x = 0 \\ 0 & si\ x \neq 0,1 \end{cases}$$

Existen familias de distribuciones que dependen de más de un parámetro, como veremos más adelante, tales como la familia binomial, la familia hipergeometrica y la familia normal.

Esperanza o valor esperado de variables aleatorias discretas:

Una universidad tiene 15000 alumnos inscriptos. Sea ***X*** la variable aleatoria:

X = "número de cursos donde se anota un alumno elegido al azar durante el año lectivo"

Se sabe que ***X*** tiene la siguiente función de probabilidad puntual:

x	1	2	3	4	5	6	7
Cantidad alumnos	150	450	1950	3750	5850	2550	300
p(*x*)	0.01	0.03	0.13	0.25	0.39	0.17	0.02

(observar que la probabilidad se obtiene dividiendo la cantidad de alumnos que toman *x* cursos por el total de alumnos de la universidad)

Si quisiéramos calcular el número promedio de cursos en los que se inscribe un alumno durante el ciclo lectivo, o valor promedio de la variable aleatoria *X*, deberíamos realizar la siguiente cuenta:

$$\text{Promedio} = \frac{1x150+2x450+3x1950+4x3750+5x5850+6x2550+7x300}{15000} =$$
$$= 1.\,p(1) + 2.p(2) + 3.p(3) + 4.p(4) + 5.p(5) + 6.p(6) + 7.p(7)$$

O sea que para calcular el promedio de la variable ***X*** alcanza con conocer su rango y las correspondientes probabilidades de cada uno de los valores que alcanza.

Sea ***X*** una variable aleatoria discreta que toma valores en R_X con función de probabilidad puntual $P_X(x)$, **la esperanza o valor esperado de la** variable ***X***, se nota E(***X***) y se define como:

$$E(\boldsymbol{X}) = \mu_X = \sum_{x_i \in R_X} x_i . p_X(x_i) \quad (\text{ siempre que } \sum_{x_i \in R_X} |x_i| . p_X(x_i) < \infty \text{ })$$

Si la suma diverge la esperanza no puede definirse y decimos que en ese caso no existe.

<u>Ejemplo 18:</u>

Hallar, si existe, la esperanza de ***X*** para cada una de las siguientes variables aleatorias:

1.- Sea *X* = "número de caras al tirar dos monedas equilibradas".

x	0	1	2
P(***X*** = *x*) = *p*(*x*)	0.25	0.5	0.25

$E(\boldsymbol{X}) = \mu_X = 0x0.25 + 1x0.5 + 2x0.25 = 1$

2.- Sea ***X*** una variable Bernoulli con: $p_X(x\,,\alpha) = \begin{cases} \alpha & si\ x = 1 \\ 1-\alpha & si\ x = 0 \\ 0 & si\ x \neq 0,1 \end{cases}$ $\alpha \in (0,1)$.

$E(\boldsymbol{X}) = \mu_X = 1.p(1) + 0.p(0) = \alpha$

3.- Sea ***X*** una variable aleatoria con la siguiente función de probabilidad puntual:

$$p_X(x) = \begin{cases} 0 & x \notin N \\ k/x^2 & x \in N \end{cases}$$ con k un número real tal que: $\sum_{i=1}^{\infty} k/i^2 = 1$

$E(\boldsymbol{X}) = \mu_X = \sum_{i=1}^{\infty} i.P(i) = \sum_{i=1}^{\infty} i.\frac{k}{i^2} = \sum_{i=1}^{\infty} \frac{k}{i} = k.\sum_{i=1}^{\infty} \frac{1}{i}$ (serie divergente)

Por lo tanto no existe la esperanza.

5.- Una ruleta tiene los números del 0 al 36. Se apuesta $1 a que sale un número par. Sea ***X*** la variable aleatoria que indica la ganancia en cada tiro.

$$\boldsymbol{X} = \begin{cases} 1 & \textit{si sale par} \\ -1 & \textit{en caso contrario} \end{cases} \qquad p(1) = \frac{18}{37} \qquad p(-1) = \frac{19}{37}$$

$E(\boldsymbol{X}) = \mu_X = 1.\frac{18}{37} + (-1)\frac{19}{37} = \frac{-1}{37} = -0.027$

Esta esperanza es negativa e indica que si jugamos un gran número de veces, por cada peso apostado esperamos perder 3 centavos aproximadamente. Recordar que si sale cero, aunque sea par gana la banca.

Interpretación de la esperanza:

- $E(\boldsymbol{X})$ es el centro de gravedad de la función de probabilidad. Es decir que si imaginamos que en cada valor x_i colocamos un masa equivalente a $p_X(x_i)$, el punto de equilibrio del sistema es $E(\boldsymbol{X})$. De esta forma $E(\boldsymbol{X})$ es una medida del "centro" de la distribución.
- La esperanza de la distribución empírica es la media aritmética que como hemos visto es el centro de gravedad de nuestra variable.
- Si imaginamos que se está desarrollando un proceso aleatorio, este resultará en un número, luego en otro, luego en otro. Uno puede preguntarse se existe algún orden en esta aleatoriedad. Los matemáticos han demostrado que en promedio los resultados de una variable aleatoria varían alrededor de la esperanza. Por ejemplo si tiramos 100 veces la moneda y registramos el número de caras, el valor esperado es 50. Sin embargo cuando tiramos varias veces 100 veces la moneda puede ser que se obtenga 57 caras, luego 46 en otra 51. Siempre los valores observados variarán alrededor de 50.

Esperanza de una función de una variable aleatoria discreta:

En una droguería se vende ácido sulfúrico en bidones de 20, 30 y 50 litros.
Sea ***X*** = "tamaño del bidón que compra el cliente"
Supongamos que la función de probabilidad puntual es:

x	20	30	50
$P(\boldsymbol{X} = x) = p(x)$	0.1	0.5	0.4

Supongamos que el precio de un bidón en función de su tamaño es:

$$f(x) = 10x - 40$$

¿Cuál es el precio esperado?

$y = f(x)$	160	260	460
P($\boldsymbol{Y} = y$) = $p(f(x))$	0.1	0.5	0.4

$E(\boldsymbol{Y}) = 160\text{x } 0.1 + 260\text{x } 0.5 + 460\text{x } 0.4 = 330$

Propiedades de la esperanza matemática:

1.- Si a y b son constantes reales cualesquiera: $E(a\,\boldsymbol{X} + b) = a.\mathrm{E}(\boldsymbol{X}) + b$.

DEMOSTRACIÓN

Sea $h(\boldsymbol{X}) = a\,\boldsymbol{X} + b$ entonces:

$$E_X(\mathrm{h}(\boldsymbol{X})) = \sum_{\substack{x\in R_X \\ x/\,p(x)>0}} h(x).p(x) = \sum_{\substack{x\in R_X \\ x/\,p(x)>0}} (ax+b).p(x) = \sum_{x\in R_X} ax.p(x) + \sum_{x\in R_X} b.p(x) =$$

$$= a.\sum_{x\in R_X} x.p(x) + b.\sum_{x\in R_X} p(x) = a.E(\boldsymbol{X}) + b$$

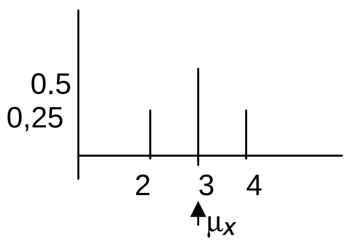
2.- Si P($\boldsymbol{X} = c$) = 1 (la variable aleatoria es constante) entonces $E(\boldsymbol{X}) = \mathrm{c}$

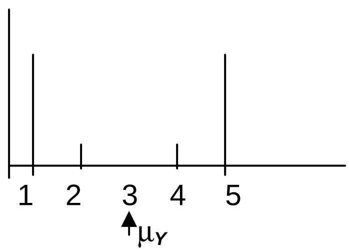
DEMOSTRACIÓN

$E(\boldsymbol{X}) = c.\,\mathrm{P}(c) = c.\,1 = c$

Varianza de una variable aleatoria discreta:

La esperanza de una variable aleatoria discreta $\boldsymbol{X}$ mide donde está centrada la distribución de probabilidad.
Hemos dicho que E($\boldsymbol{X}$) es el punto de equilibrio de la función de probabilidad de la variable.
Consideremos dos variables aleatorias $\boldsymbol{X}$ e $\boldsymbol{Y}$ con las siguientes probabilidades puntuales:





Queremos medir la variabilidad de las variables $\boldsymbol{X}$ e $\boldsymbol{Y}$. Una forma de hacerlo es promediando el cuadrado de los desvíos respecto a la esperanza, de igual forma que calculábamos la variabilidad de una muestra promediando el cuadrado de los desvíos respecto del promedio muestral.

Sea $\boldsymbol{X}$ una variable discreta con función de probabilidad $p_X(x)$ y esperanza igual a μ.

a) La varianza de la variable ***X*** se nota con $V(\boldsymbol{X}) = \sigma^2 = \sigma_X^2$ se define como:

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

b) El desvío estándar se define como: $\sigma_X = \sqrt{V(X)}$

Observemos que si la mayor parte de la distribución de los valores de ***X*** está cerca de μ entonces σ^2 es pequeña, pero si hay valores de ***X*** muy alejados de μ con $p_X(x)$ grande, entonces σ^2 es grande.

<u>Ejemplo 19:</u>

1.- Sea ***X*** = “número de caras al tirar dos monedas”

x	0	1	2
P(X = x) = p(x)	0.25	0.5	0.25

$E(\boldsymbol{X}) = \mu_X$ = 0x0.25 + 1x0.5 + 2x0.25 = 1

$V(\boldsymbol{X}) = E\left((X-1)^2\right) = (0-1)^2\ p(0) + (1-1)^2\ p(1) + (2-1)^2\ p(2) = 0.25 + 0.25 = 0.50$

2.- Sea ***X*** una variable Bernoulli con: $P_X(x, \alpha) = \begin{cases} \alpha & si\ x=1 \\ 1-\alpha & si\ x=0 \\ 0 & si\ x \neq 0,1 \end{cases}$ $\alpha \in (0,1)$.

$E(\boldsymbol{X}) = \mu_X = 1.p(1) + 0.p(0) = \alpha$

$V(\boldsymbol{X}) = E\left((X-\alpha)^2\right) = (1-\alpha)^2\ p(1) + (0-\alpha)^2\ p(0) = (1-\alpha)^2\ \alpha + \alpha^2\ (1-\alpha) =$

$= \alpha\ (1-\alpha)\ (1-\alpha+\alpha) = \alpha\ (1-\alpha)$

Resultado importante:

Si ***X*** es Bernoulli de parámetro α, entonces $E(\boldsymbol{X}) = \alpha$ y $V(\boldsymbol{X}) = \alpha\ (1-\alpha)$.

Proposición:

$$V(\boldsymbol{X}) = E\left(X^2\right) - \left(E(X)\right)^2 = E\left(X^2\right) - \mu_X^2$$

DEMOSTRACIÓN

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

$$= \sum_{x \in R_X} x^2.p(x) - \sum_{x \in R_X} 2\mu_X x.p(x) + \sum_{x \in R_X} \mu_X^2.p(x)$$

$$= \sum_{x \in R_X} x^2.p(x) - 2\mu_X.\sum_{x \in R_X} x.p(x) + \mu_X^2.\sum_{x \in R_X} p(x) =$$

$$= \sum_{x \in R_X} x^2.p(x) - 2\mu_X.E(X) + \mu_X^2.1 = \sum_{x \in R_X} x^2.p(x) - 2\mu_X^2 + \mu_X^2.1 =$$

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

Propiedades de la varianza y el desvío estándar:

1.- Para cualquier función $h(\boldsymbol{X})$ se verifica: $V(h(\boldsymbol{X})) = \sum_{x \in R_X} \left(h(x) - E(h(X))\right)^2 p(x)$

2.- Si $h(\boldsymbol{X}) = a\boldsymbol{X} + b \Rightarrow V(h(\boldsymbol{X})) = V(a\boldsymbol{X} + b) = a^2.V(\boldsymbol{X})$ y $\sigma_{aX+b} = |a| \,.\sigma_X$

3.- En particular: $\sigma^2_{aX} = a^2 \,.\sigma^2_X$ *cambios de escala afectan la varianza*

$\sigma^2_{X+b} = \sigma^2_X$ *una traslación no afecta la varianza*

Demostración de 2:

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

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

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

En particular si a=0, es decir Y=b con probabilidad 1 entonces $V(Y)=a^2V(X)=0$

Algunas distribuciones importantes

A. Distribución binomial

Muchos experimentos aleatorios satisfacen las siguientes condiciones, a veces en forma exacta y otras en forma aproximada:

1. El experimento consiste en realizar *n* pruebas, con *n* un número fijo.
2. Las pruebas son todas idénticas y cada prueba tiene sólo dos resultados (siempre los mismos) que denominaremos éxito o fracaso.
3. Las pruebas son independientes, es decir el resultado de una prueba no influye sobre el resultado de cualquier otra.
4. La probabilidad *p* de éxito se mantiene constante a traves de los ensayos.

Un experimento que satisface las cuatro condiciones anteriores se denomina ***experimento binomial*** y cada prueba se denomina ***ensayo Bernoulli***.

Ejemplos:

1.- Tiramos una moneda *n* veces y llamamos éxito a que salga cara.

2.- Una fábrica produce tubos fluorescentes. Se toma una muestra de 50 tubos al azar y se los enciende. Se considera Éxito: “el tubo dura más de 400 horas”

3.- Se extraen 3 bolillas sin reposición de una caja que contiene 5 rojas y 5 blancas. Se considera Éxito: “la bolilla extraída es roja”. ¿Es binomial?

No, pues la P(de sacar la un roja dado que en el ensayo anterior salió roja) varía de un ensayo a otro ensayo. Los ensayos no son independientes.

4.- Una fábrica produce 50000 arandelas por día. Supongamos que de ellas 45000 son "buenas", en el sentido que cumplen con las normas IRAM. Se extrae una muestra aleatoria de 10 arandelas, obviamente la extracción es sin reposición. Se considera como Éxito: "la arandela cumple con las normas IRAM" ¿Es binomial?

Aunque al igual que en el caso anterior las repeticiones no son independientes, el resultado de una extracción afecta poco al resultado de la siguiente. Esto ocurre pues el tamaño de la muestra es mucho menor que el de la población (10<<50000)

$$P\left(E_2 / E_1\right) = \frac{44999}{49999} \cong 0.9$$

$$P\left(E_{10} / E_1 \cap E_2 \cap ... \cap E_9\right) = \frac{44991}{49991} \cong 0.9$$

Aún cuando las pruebas no son independientes, la probabilidad de Éxito cambia muy poco de un experimento a otro. Por lo tanto podemos aproximar el modelo por un modelo binomial.

Observación:

La aproximación será razonable siempre que $\frac{n}{N} \leq 0.05$ y *p* no se encuentre muy cerca de 0 ni de 1. Esto equivale a decir que *n* es a lo sumo el 5% del tamaño de la población.

Variable aleatoria binomial:

Consideremos un experimento binomial que consiste de *n* ensayos y sea *p* la probabilidad de éxito. Denominamos variable aleatoria binomial de parámetros *n* y *p* a la variable aleatoria ***X*** que cuenta el número de éxitos en las *n* pruebas.

Notación: $X \sim B_i(n, p)$ *n* = número de ensayos *p* = probabilidad de éxito

Calculemos la función de probabilidad puntual de ***X***:

$R_X = \{0, 1, 2, ..., n\}$

Sea *k* un entero tal que: $0 \leq k \leq n$

Un resultado posible con *k* éxitos y por lo tanto $n - k$ fracasos es:

$\underbrace{EE... E}_{k}\ \underbrace{FF... F}_{n-k} \Rightarrow$ P(EE.. E FF....F) = $p^k (1- p)^{n-k}$

¿Cuántas sucesiones de *k* éxitos y *n-k* fracasos podemos encontrar?

Encontraremos tantas distintas como cantidad de formas de seleccionar entre los *n* lugares disponibles (cantidad de ensayos) los *k* lugares donde ubicar los éxitos. Esto se obtiene con: $\binom{n}{k}$

Por lo tanto: $p_X(k) = P(X = k) = \binom{n}{k}.p^k.(1-p)^{n-k}$

Comprobemos que: $\sum_{k=0}^{n} p_X(k) = 1$

$$\sum_{k=0}^{n} p_X(k) = \sum_{k=0}^{n}\binom{n}{k}.p^k.(1-p)^{n-k} = \left(p+(1-p)\right)^n = 1^n = 1$$

Observación:

El binomio de Newton establece que: $(a+b)^n = \sum_{j=0}^{n} \binom{n}{k} . a^j . b^{n-j}$

Notación:

$b(k,n,p)$ = P(***X*** = *k*) si $X \sim B_i(n,p)$

Ejemplo 20:

Un alumno especula con aprobar un examen de biología sin saber absolutamente nada. El examen es tipo de elección múltiple: consta de 15 preguntas, cada una con 5 opciones de respuesta y 1 sola respuesta correcta en cada pregunta. Si el alumno elige en forma absolutamente aleatoria sus respuestas y el examen se considera aprobado con 8 o más preguntas bien contestadas:

a) ¿cuál es la probabilidad de que conteste bien a lo sumo 7 preguntas?
b) ¿cuál es la probabilidad de que el alumno apruebe el examen?
c) ¿cuál es la probabilidad de que acierte entre 4 y 7 respuestas?
d) ¿cuál es la probabilidad de que conteste exactamente 7 preguntas bien?

Observar que la variable ***X*** = "número de respuestas correctas en el examen" es una variable binomial con *n* = 15 y $p = \frac{1}{5} = 0.2 \Rightarrow X \sim B_i(15, 0.2)$

a) $P(X \le 7) = B(7,15,0.2) = \sum_{k=0}^{7} \binom{15}{k} 0.2^k x\, 0.8^{n-k}$ = 0.996 (tabla)

b) $P(X \ge 8)$ = 1 - $P(X \le 7)$ = 1 – 0.996 = 0.004

c) $P(4 \le X \le 7)$ = F(7) – F(4^-) = $B(7,15,0.2)$ - $B(3,15,0.2)$ = 0.996 – 0.648 = 0.348

d) $p_X(7)$ = $P(X = 7)$ = F(7) – F(7^-) = F(7) – F(6) = 0.996 – 0.982 = 0.014

Esperanza y varianza de $X \sim B_i(n,p)$:

Recordemos que si ***X*** es Bernoulli con probabilidad de éxito *p*, entonces *E*(***X***) = *p* y V(***X***) = *p*(1-*p*)

Observar que si ***X*** es Bernoulli con probabilidad de éxito *p*, entonces $X \sim B_i(1,p)$.

Proposición:

Si $X \sim B_i(n,p)$ entonces: *E*(***X***) = *n.p* y *V*(***X***) = *n.p.*(1-*p*)

Dem:

$$E(X) = \sum_{k=0}^{n} k.p_X(k) = \sum_{k=0}^{n} k.\binom{n}{k}.p^k.(1-p)^{n-k} = 0 + \sum_{k=1}^{n} k.\binom{n}{k}.p^k.(1-p)^{n-k} =$$

$$= \sum_{k=1}^{n} k.\frac{n!}{k!(n-k)!}.p^k.(1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!}.p^k.(1-p)^{n-k} =$$

$$= n.p.\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!}.p^{k-1}.(1-p)^{(n-1)-(k-1)}$$

$$= n.p.\sum_{i=0}^{n-1} \frac{(n-1)!}{i!(n-1-i)!}.p^i.(1-p)^{(n-1)-i} = n.p(p+(1-p))^{n-1} = n.p$$

$$V(\boldsymbol{X}) = E\left(X^2\right)-\left[E(X)\right]^2 = E\left(X^2\right)-n^2.p^2$$

$$E\left(X^2\right) = \sum_{k=0}^{n} k^2.\binom{n}{k}.p^k.(1-p)^{n-k} = \sum_{k=0}^{n} k((k-1)+1).\binom{n}{k}.p^k.(1-p)^{n-k} =$$

$$= \sum_{k=0}^{n} k(k-1).\binom{n}{k}.p^k.(1-p)^{n-k} + \sum_{k=0}^{n} k.\binom{n}{k}.p^k.(1-p)^{n-k} =$$

$$= \sum_{k=0}^{n} k(k-1).\frac{n!}{k!(n-k)!}.p^k.(1-p)^{n-k} + E(X) =$$

$$= 0 + 0 + \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!}.p^k.(1-p)^{n-k} + n.p =$$

$$= n.(n-1).p^2.\sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-k)!}.p^{k-2}.(1-p)^{n-k} + n.p =$$

$$= n.(n-1).p^2.\sum_{j=0}^{n-2} \frac{(n-2)!}{j!(n-2-j)!}.p^{j}.(1-p)^{n-2-j} + n.p =$$

$$= n.(n-1).p^2.1 + n.p = n.p.\left[(n-1).p+1\right]$$

$$V(\boldsymbol{X}) = E\left(X^2\right)-\left[E(X)\right]^2 = E\left(X^2\right)-n^2.p^2 = n.p.\left[(n-1).p+1\right] - n^2.p^2 =$$

$$= n.p(n.p-p+1-n.p) = n.p.(1-p)$$

B. Distribución hipergeométrica:

Supongamos que:

- La población muestreada consta de **N** individuos (población finita).
- Cada individuo puede ser éxito fracaso y hay **D** éxitos en la población.
- Se extrae una muestra de n individuos, de forma tal que cualquier subconjunto de tamaño *n* tiene igual probabilidad de ser elegido.

Sea ***X*** = "número de éxitos en la muestra". Esta variable depende de: *n*, **N** y **D**.

Notación: $X \sim H(n,N,D)$ ***X***: tiene una distribución hipergeométrica

X equivale a una binomial pero sin reemplazo (sin reposición).

Proposición:

Si $X \sim H(n,N,D) \Rightarrow p(x) = \begin{cases} \dfrac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} & si \ \max(0,n-(N-D)) \le x \le min(n,D) \\ 0 & en \ \ otro \ \ caso \end{cases}$

Notación:

$$X \sim H(n,N,D) \Rightarrow p(x) = h(x,n,N,D)$$

Dem:

La cantidad de subconjuntos de n elementos es: $\binom{N}{n}$

La cantidad con *n* elementos y *x* éxitos (*n* - *x* fracasos) es: $\binom{D}{x}\binom{N-D}{n-x}$

Observar que: 1.- $x \leq$ **D** ^ $x \leq n \Rightarrow x \leq$ mín(**D**, *n*)

2.- $x \geq 0$ ^ $n - x \leq$ **N** – **D** $\Rightarrow x \geq 0$ ^ $x \geq n$ – (**N** – **D**) $\Rightarrow$ $x \geq$ max (0, *n* – (**N** – **D**))

3.- $\sum p(x) = 1$

Proposición:

Si $X \sim H(n,N,D) \Rightarrow E(X) = n.\frac{D}{N}$ y $V(X) = \frac{N-n}{N-1}.n.\frac{D}{N}\left(1-\frac{D}{N}\right)$

<u>*Observación*</u>:

Si llamamos $p = \frac{D}{N} \Rightarrow E(X) = n.p$ y $V(X) = \frac{N-n}{N-1}.n.p.(1-p)$

Al factor $\frac{N-n}{N-1}$ se lo llama factor de corrección para población finita. Si *n* es muy pequeño con relación a ***N***, entonces el factor de corrección es aproximadamente 1, o sea se comporta como en el caso de muestras con reposición o población infinita.

Aproximación de la hipergeométrica por la binomial:

Ya hemos dicho que si $\frac{n}{N} << 0.05$ entonces la binomial es una buena aproximación a la hipergeométrica.

Una regla más precisa es la siguiente:

Sean **N** = tamaño de la población y **D** = número de éxitos; dos cantidades tales que si **N** crece **D** también crece pero $\frac{D}{N}$ es fijo. Entonces $h(x,n,N,D)$ puede aproximarse por $b(x,n,p)$ siempre que $p = \frac{D}{N}$ no esté muy cerca de 0 ó de 1 y el cociente $\frac{n}{N}$ sea pequeño.

C. ***Distribución de Poisson***:

No existe un experimento simple sobre el cual se base la distribución de Poisson, como en el caso de la binomial o la hipergeometrica, aunque puede pensarse como la distribución límite de una variable binomial bajo ciertas condiciones.

Definición:

Una variable aleatoria ***X*** tiene distribución de Poisson de parámetro λ, si su función de probabilidad puntual es:

$$p_X(x) = \begin{cases} \dfrac{e^{-\lambda}.\lambda^x}{x!} & x \in N_0 \\ 0 & x \notin N_0 \end{cases} \qquad \text{con} \quad \lambda > 0$$

El parámetro λ frecuentemente representa una tasa de ocurrencia o intensidad del proceso de algún fenómeno por unidad de área ó unidad de tiempo.

Notación: $X \sim P(\lambda)$

Observación:

1.- si $\lambda > 0 \Rightarrow p_X(x) \geq 0$

2.- $\sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} e^{-\lambda}.\dfrac{\lambda^k}{k!} = e^{-\lambda}.\sum_{k=0}^{\infty}\dfrac{\lambda^k}{k!} = e^{-\lambda}.e^{\lambda} = 1$

(recordar la expansión en serie de Mac Laurin de e^x)

3.- $e^x = 1+\dfrac{x}{1!}+\dfrac{x^2}{2!}+...+\dfrac{x^n}{n!}+....$

Ejemplo 21:

Sea ***X*** = "número de árboles infectados por una enfermedad en una hectárea de bosque". Supongamos X~ P(5) donde λ=5 es la tasa de infección por hectárea

Calcular la probabilidad de que haya exactamente dos árboles infectados

$$p_X(2) = \frac{e^{-5}.5^2}{2!} = 0.084$$

Calcular la probabilidad de que haya a lo sumo dos árboles infectados

$$P(\boldsymbol{X} \leq 2) = \sum_{k=0}^{2}\frac{e^{-5}.5^k}{k!} = e^{-5}\left(1+\frac{5^1}{1!}+\frac{5^2}{2!}\right) = 0.125$$

Proposición:

Si $X \sim P(\lambda) \Rightarrow E(X) = \lambda$ y $V(X) = \lambda$

DEMOSTRACIÓN

$$E(X) = \sum_{x=0}^{\infty} x.e^{-\lambda}.\frac{\lambda^x}{x!} = 0 + \sum_{x=1}^{\infty} x.e^{-\lambda}.\frac{\lambda^x}{x!} = \lambda.\sum_{x=1}^{\infty} e^{-\lambda}.\frac{\lambda^{x-1}}{(x-1)!} = \lambda.\sum_{j=0}^{\infty} e^{-\lambda}.\frac{\lambda^j}{j!} = \lambda.1 = \lambda$$

$$V(X) = E\left(X^2\right) - \left(E(X)\right)^2 = E\left(X^2\right) - \lambda^2$$

$$E\left(X^2\right) = \sum_{x=0}^{\infty} x^2.e^{-\lambda}.\frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} x(x-1).e^{-\lambda}.\frac{\lambda^x}{x!} + \sum_{x=0}^{\infty} x.e^{-\lambda}.\frac{\lambda^x}{x!} = \sum_{x=2}^{\infty} x(x-1).e^{-\lambda}.\frac{\lambda^x}{x!} + E(X)$$

$$= \sum_{x=2}^{\infty} e^{-\lambda}.\frac{\lambda^2.\lambda^{x-2}}{(x-2)!} + \lambda = \lambda^2.\sum_{k=0}^{\infty} e^{-\lambda}.\frac{\lambda^k}{k!} + \lambda = \lambda^2 + \lambda$$

$$V(X) = E\left(X^2\right) - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

La distribución de Poisson como distribución límite de la binomial:

La distribución de Poisson puede obtenerse como límite de la binomial haciendo tender *n* a infinito con *p* tendiendo a cero de modo que *n.p* = cte.

Proposición:
Sea $X \sim Bi(n, p)$. Supongamos que: $n \to \infty$ y $p \to 0$ de forma tal que $n.p = \lambda$ (fijo),
entonces: $p(x) \to \dfrac{e^{-\lambda}\lambda^{x}}{x!}$ para todo *x* natural o cero.

<u>Ejemplo 22:</u>

Se arrojan dos dados 100 veces, se cuenta el número de veces que sale doble seis.

$X \sim Bi\left(100, \dfrac{1}{36}\right)$ $n.p = 2.78 \leq 20$ $p = 0.0278 > 0.01$

k	*Binomial*	*Poisson*
0	0.0596	0.0620
1	0.1705	0.1725
2	0.2414	0.2397
5	0.0858	0.0858
8	0.0050	0.0055
9	0.0015	0.0017
10	0.0004	0.0005

Como se puede observar la aproximación es bastante buena aunque no se cumplen todos los requisitos de la regla.

<u>Ejemplo 23:</u>

Supongamos que el número de llamadas que se reciben en una oficina sigue un proceso Poisson de parámetro v = 0,5 llamadas por minuto.
¿Cuál es la probabilidad de que no se reciban llamadas entre las 12 hs y las 12hs 5 m?

X_{5m} = "número de llamadas en 5 minutos"

X_{5min} ~P(2,5) con $\lambda = v.t$ = 0.5 x 5 = 2.5

P(X_{5m} = 0) = $\dfrac{e^{-\lambda}.\lambda^{0}}{0!} = e^{-2.5}$ = 0.082

¿Cuál es el número esperado de llamadas en 1 hora?

X_{60m} = "número de llamadas en 60 minutos"

X_{60min} ~P(30) con $\lambda = v.t$ = 0.5 x 60 = 30

$E\left(X_{60m}\right) = \lambda = v.t = 30$

¿Cuál es la probabilidad de que no se reciban llamadas entre las 16 hs y las 16hs 5 m?
Idem que a), discutir si es razonable suponer que la variable es Poisson.

Nota:

Existen variables generadas por un proceso de Poisson que no están relacionadas con el tiempo, sino que ocurren en 2 o 3 dimensiones. Ejemplos:

- número de fallas en una tira de papel
- número de árboles infectados por hectárea
- número de accidentes a lo largo de la ruta
- número de fallas por unidad de volumen de cemento para hormigón en una construcción

En estos casos v = “número de fallas por unidad de área o de volumen”

Variables aleatorias continuas

Una variable aleatoria se dice continua, si el conjunto de valores posibles de la variable es un intervalo, es decir existen a y b tales que a < b para los cuales se cumple que x es un valor posible de la variable para todo x $\in$ (a, b). En particular este intervalo puede ser de la forma (a, ∞), o ($-\infty$,b) o ($-\infty$, ∞).

Por ejemplo:

1.- Se mide la profundidad de un lago en diferentes puntos elegidos al azar.

2.- Se elige un número al azar entre el 0 y el 1.

3.- Se mide la altura de un adulto seleccionado al azar de cierta población.

Se dice que ***X*** es una **variable aleatoria continua** si existe una función de densidad o distribución de probabilidad de ***X***, notada f_X tal que para todo *a* y *b*, tales que $a \le b$:

$$P(a \le x \le b) = \int_a^b f_X(x)\,dx$$

Para que f_X sea una función de densidad, debe cumplir las siguientes dos condiciones simultáneamente:

1.- $f_X(x) \ge 0 \quad \forall\ x$

2.- $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

Función indicadora sobre un intervalo:

$$I_{[a,b]}(x) = \begin{cases} 0 & si\ x < a \\ 1 & si\ \ a \le x \le b \\ 0 & si\ x > b \end{cases}$$

Ejemplo 24:

Sea $f(x) = a.x^2.I_{[1,3]}(x)$

a) ¿cuánto debe valer a para que f resulte una función de densidad?

$$1 = \int_{-\infty}^{+\infty} f_X(x)\,dx = \int_{-\infty}^{+\infty} a.x^2.I_{[1,3]}(x)\,dx = \int_{1}^{3} a.x^2\,dx = a.\left.\frac{x^3}{3}\right|_1^3 = a.\left(3^2 - \frac{1}{3}\right) = a.\frac{26}{3}$$

$$1 = a.\frac{26}{3} \Rightarrow a = \frac{3}{26}$$

b) Calcular P($X \geq 2$)

$$P(X \geq 2) = \int_{2}^{+\infty} f_X(x)\,dx = \int_{2}^{+\infty} a.x^2.I_{[1,3]}(x)\,dx = \int_{2}^{3} \frac{3}{26}.x^2\,dx = \frac{3}{26}\left.\frac{x^3}{3}\right|_2^3 = \frac{1}{26}\left(3^3 - 2^3\right) = \frac{19}{26}$$

Sea ***X*** una variable aleatoria continua con función de densidad f_X entonces **la función de distribución de probabilidades acumulada** de ***X***, F_X , se define como:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\,dt \qquad \forall\ x \in \mathbf{R}$$

En muchos casos puede ser de utilidad definir la **función de distribución de probabilidades acumuladas a derecha.**

$$G_X(x) = P(X > x) = \int_{x}^{\infty} f_X(t)\,dt \qquad \forall\ x \in \mathbf{R}$$

Se satisface claramente a partir de esta definición que $\forall\ x \in$ R:

$$\mathbf{F_X(x) + G_X(x) = 1}$$

Proposición:

Sea ***X*** una variable aleatoria continua con función de densidad f_X y función de distribución acumulada F_X , entonces $\forall\ a$, $\forall\ b$ tales que $a < b$ se verifica que:

i.- $P(a \leq x \leq b) = F_X(b) - F_X(a)$

ii.- $P(X = a) = F_X(a) - F_X(a) = 0$

Propiedad:

Sea ***X*** una variable aleatoria continua con función de densidad f_X y función de distribución acumulada F_X , entonces en todo punto x donde F_X es derivable:

$$F'_X(x) = f_X(x)$$

Esta demostración es inmediata aplicando el Teorema Fundamental del Cálculo Integral.

Algunos ejemplos de funciones de distribución con nombre propio:

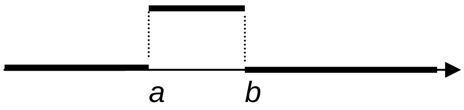
D. Distribución uniforme:

Se dice que ***X*** tiene una distribución uniforme en el intervalo $[a, b]$ si su función de densidad es:

$$f_X(x) = \frac{1}{b-a}.I_{[a,b]}(x)$$

Notación:

$$X \sim U[a,\ b]$$



La función de distribución acumulada es:

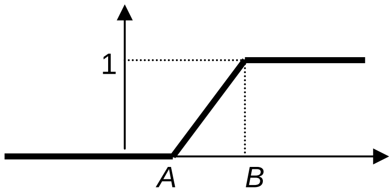
Si $x < a \quad \Rightarrow \quad F_X(x) = 0$

Si $a \leq x < b \quad \Rightarrow \quad F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{a}^{x} \frac{1}{b-a}\,dt = \left.\frac{t}{b-a}\right|_{a}^{x} = \frac{x-a}{b-a}$

Si $x \geq b \quad \Rightarrow \quad F_X(x) = 1$

Por lo tanto:

$$F_X(x) = \begin{cases} 0 & si\ x < a \\ \dfrac{x-a}{b-a} & si\ a \leq x < b \\ 1 & si\ x \geq b \end{cases}$$



Ejemplo 25:

Los colectivos llegan a una parada cada 15 minutos exactamente. Supongamos que un pasajero llega a la parada en un instante uniformemente distribuido entre las 7:00 y las 7:30hs.

Hallar la probabilidad de que espere:

a- menos de 5 minutos

b- más de 10 minutos

a- P(7:10<X<7:15)+ P(7:25<X<7:30)= 5/30+5/30=1/3

b- P(7<X<7:05)+ P(7:15<X<7:20)= 5/30+5/30=1/3

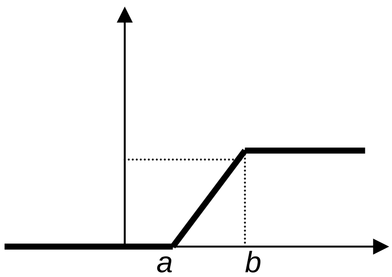
OBSERVACIÓN: en esta variable aleatoria la probailidad del intervalo (xo,xo+h) sólo depende de h.

Veamos que esto se observa en la función de distribución de probabilidades acumuladas.

Esta funcion tiene un crecimiento lineal.

En cada subintervalo (xo, xo+h), se cumple

Que F(xo+h)-F(xo)= h/(b-a)

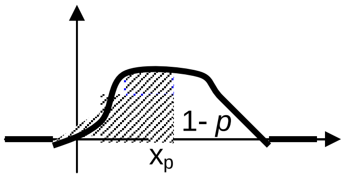


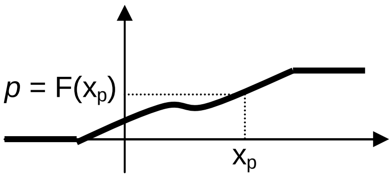
Percentiles de una distribución continua:

Sea $0 < p < 1$. El percentil (100*p* - ésimo de la distribución de una variable aleatoria continua *X*, con función de densidad f_X, se define como el valor x_p de la variable tal que:

$$F_X(x_p) = p = \int_{-\infty}^{x_p} f_X(t)\,dt$$

O sea que x_p es el valor de la variable que deja a la izquierda un área p, y a la derecha un área $1-p$.





Ejemplos:

1.- En toda variable aleatoria continua la mediana de la variable es $\tilde{\mu} = x_{0.5}$
En los gráficos de las distribuciones de probabilidad del ejemplo anterior se ha marcado con color rojo la mediana de cada una de las distribuciones. Veamos que diferentes son… cuál es el motivo?

2.- Para la variable con función de densidad $f_X(x) = \dfrac{x}{2} I_{[0,2]}(x)$, obtener:

a.- la función de distribución acumulada

si x < 0 $\Rightarrow F_X(x) = 0$

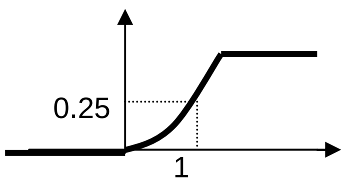
si $0 \le x < 2 \Rightarrow F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{x} \frac{t}{2} I_{[0,2]}(t)\,dt = \int_{0}^{x} \frac{t}{2}\,dt = \left.\frac{t^2}{4}\right|_0^x = \frac{x^2}{4}$

si x ≥ 2 $\Rightarrow F_X(x) = 1$

b.- el percentil 25 (primer cuartil).

Buscamos $0 < x_{0.25} < 2$ tal que $F_X(x_{0.25}) = 0.25 \Rightarrow$

$\dfrac{(x_{0.25})^2}{4} = 0.25 \Rightarrow (x_{0.25})^2 = 1 \Rightarrow (x_{0.25}) = 1$



Esperanza de una variable aleatoria continua:

Sea ***X*** una variable aleatoria continua con función de densidad f_X entonces:

$E(X) = \mu_X = \int_{-\infty}^{+\infty} x.f_X(x)\,dx$ siempre que: $\int_{-\infty}^{+\infty} |x|.f_X(x)\,dx < \infty$

La integral debe ser finita para que afirmemos que existe la esperanza de ***X***.

Ejemplo:

Si $X \sim U[a,b]$ calcular $E(X)$

$$\mu_X = \int_{-\infty}^{+\infty} x.f_X(x)\,dx = \int_{-\infty}^{+\infty} x\frac{1}{b-a} I_{[a,b]}(x)\,dx = \int_{a}^{b} x.\frac{1}{b-a}\,dx = \left.\frac{t^2}{2(b-a)}\right|_a^b =$$

$$= \frac{b^2-a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)} = \frac{b+a}{2}$$

Varianza de una variable aleatoria continua:

Sea ***X*** una variable aleatoria continua con función de densidad f_X entonces

1.- la varianza de X es: $V(X) = \sigma_X^2 = E\left((X-\mu_X)^2\right) = \int_{-\infty}^{+\infty} (x-\mu_X)^2 . f_X(x)\,dx$

2.- el desvío estándar de X es: $\sigma_X = \sqrt{V(X)}$

Proposición:

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

2.- $V(aX+b) = a^2.V(X)$

Dem:

1.- $V(X) = E\left((X-\mu_X)^2\right) = \int_{-\infty}^{+\infty} (x-\mu_X)^2 . f_X(x)\,dx = \int_{-\infty}^{+\infty} (x^2 - 2x\mu_X + {\mu_X}^2). f_X(x)\,dx =$

$$= \int_{-\infty}^{+\infty} x^2 . f_X(x)\,dx + \int_{-\infty}^{+\infty} (-2x\mu_X). f_X(x)\,dx + \int_{-\infty}^{+\infty} {\mu_X}^2 . f_X(x)\,dx =$$

$$= \int_{-\infty}^{+\infty} x^2 . f_X(x)\,dx + (-2\mu_X)\int_{-\infty}^{+\infty} x . f_X(x)\,dx + {\mu_X}^2 . \int_{-\infty}^{+\infty} f_X(x)\,dx =$$

$$= \int_{-\infty}^{+\infty} x^2 . f_X(x)\,dx - 2\mu_X E(X) + {\mu_X}^2 .1 = E\left(X^2\right) - {\mu_X}^2 = E\left(X^2\right) - E^2(X)$$

2.- $V(aX+b) = \int_{-\infty}^{+\infty} (ax+b-\mu_{aX+b})^2 . f_X(x)\,dx = \int_{-\infty}^{+\infty} (ax+b-(a\mu_X+b))^2 . f_X(x)\,dx =$

$$= \int_{-\infty}^{+\infty} a^2 (x-\mu_X)^2 . f_X(x)\,dx = a^2 . \int_{-\infty}^{+\infty} (x-\mu_X)^2 . f_X(x)\,dx = a^2.V(X)$$

<u>Ejemplo 26:</u>

Sea $X \sim U[a,b]$ calcular $V(X)$.

$$V(X) = E\left(X^2\right) - E^2(X) = = \int_{-\infty}^{+\infty} x \frac{1}{b-a} I_{[a,b]}(x)\,dx - \left(\frac{a+b}{2}\right)^2 =$$

$$V(X) = \int_a^b x^2 . \frac{1}{b-a}\,dx - \left(\frac{a+b}{2}\right)^2 = \left.\frac{x^3}{3(b-a)}\right|_a^b - \left(\frac{a+b}{2}\right)^2 =$$

$$= \frac{b^3-a^3}{3(b-a)} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} - \frac{a^2+2ab+b^2}{4} =$$

$$= \frac{4(b^2+ab+a^2)}{12} - \frac{3(a^2+2ab+b^2)}{12} = \frac{(a^2-2ab+b^2)}{12} = \frac{(b-a)^2}{12}$$

Propiedad:

Si $X \sim U\left[a,b\right] \Rightarrow E(X) = \dfrac{b+a}{2}$ y $V(X) = \dfrac{(b-a)^2}{12}$

E. Distribución normal

La variable aleatoria ***X*** tiene distribución normal de parámetros μ y σ^2 (con $\mu \in$ **R** y $\sigma > 0$) si su función de densidad es:

$$f_X(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) \qquad -\infty < x < +\infty$$

Notación: $X \sim N(\mu,\sigma^2)$

Veamos que $f_X(x,\mu,\sigma)$ cumple con las propiedades de una función de densidad:

1.- $f_X(x,\mu,\sigma) \geq 0$ pues : $\dfrac{1}{\sqrt{2\pi\sigma^2}} > 0$ y $e^{-\frac{1}{2\sigma^2}(x-\mu)^2} > 0$ cualquiera sea *x* real.

2.- $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} = 1$

La demostración de este hecho requiere manejo de integrales dobles, por eso se omite.

<u>Nota</u>:
La normal fue introducida por el matemático Abraham De Moivre en 1733, quien la usó para aproximar probabilidades binomiales cuando el parámetro *n* es grande. Este resultado fue extendido luego por Laplace y es lo que hoy se conoce como teorema central del límite.

Observaciones:

- El gráfico de la función de densidad de la normal tiene forma de campana con eje de simetría la recta $x = \mu$; y los puntos de inflexión en $x = \mu - \sigma$ y $x = \mu + \sigma$.
- La gráfica de la función normal es asintótica respecto del eje de absisas.
- Los parámetros de la normal indican: μ la posición y σ la dispersión.
- La distribución normal es importante no sólo porque hay variables naturales que siguen esta distribución (por ejemplo los errores aleatorios en el proceso de medición de magnitudes) sino porque bajo ciertas condiciones suele ser una buena aproximación a la distribución de otras variables

Distribución normal estándar:
La distribución normal con parámetros $\mu = 0$ y $\sigma = 1$ se denomina distribución normal estándar y se nota Z~N(0;1) . La función de distribución de una normal estándar se nota $\phi(z)$.

Por lo tanto: si Z~N(0;1) $\Rightarrow$ $f(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} = \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}z^2\right)$ -∞ < z < +∞

$$F(z) = \phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt$$

La función de distribución de la normal estándar está tabulada y también disponible en Excel:

Uso de tablas:

1.- $P(z \le 1.25) = \phi(1.25)$ = 0.8944

2.- $P(z > 1.25)$ = 1 - $\phi(1.25)$ = 0.1056

3.- $P(-0.38 \le z \le 1.25)$ = $\phi(1.25)$ - $\phi(-0.38)$ = 0.8944 – 0.3520 = 0.5424

Percentiles de la distribución normal estándar:

Sea 0 < *p* < 1 el percentil (100*p*) ó percentil (100*p*)-avo es el valor de *Z* (z_{1-p}) tal que $\phi(z_{1-p})$ = *p*.

Ejemplo:
Calcular el percentil 99 de la distribución normal estándar.

$\phi(z_{0.01})$ = 0.99 $\Rightarrow$ z_1 = 2.32 corresponde a p_1 = 0.9898

z_2 = 2.33 corresponde a p_2 = 0.9901

Para calcular $z_{0.01}$ se realiza una interpolación lineal entre los dos valores de z encontrados: $z_{0.01} = 2.32 + \dfrac{0.99 - 0.9898}{0.9901 - 0.9898}\, 0.01 = 2.32 + \dfrac{0.0002}{0.0003}\, 0.01$

= 2.32 + 0.007 = 2.327

Algunos prefieren elegir el más próximo a 0.99 en este caso 0.9901 diciendo entonces que el precentil 99 es $z_{0.01}$ = 2.33 .

Notación:
Llamaremos z_α al valor de la variable $N(0,1)$ que deja a su derecha un área de α.

Ejemplo 27:

Si $X \sim N(\mu, \sigma^2)$ y $Z \sim N(0,1) \Rightarrow$

1.- $P(|z| \le 1)$ = $\phi(1)$ - $\phi(-1)$ = 0.6826

2.- $P(|x-\mu| \le \sigma) = P(-\sigma \le x-\mu \le \sigma) = P\left(-1 \le \dfrac{x-\mu}{\sigma} \le 1\right) = P(-1 \le z \le 1)$ = 0.6826

3.- $P(|x-\mu| \le 2\sigma) = P(|z| \le 2)$ = 0.9544

4.- $P\left(|x-\mu|\leq 3\sigma\right) = P\left(|z|\leq 3\right)$ = 0.9974

Proposición:

1.- Si $Z \sim N(0,1) \Rightarrow E(Z) = 0$ y $V(Z) = 1$

$f_Z(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right) \Rightarrow g(z) = z.\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)$ es una función impar, por lo tanto:

$$E(Z) = \int_{-\infty}^{+\infty} g(z)dz = \int_{-\infty}^{0} g(z)dz + \int_{0}^{+\infty} g(z)dz = -\int_{+\infty}^{0} g(-z)dz + \int_{0}^{+\infty} g(z)dz =$$

$$= \int_{0}^{+\infty} g(-z)dz + \int_{0}^{+\infty} g(z)dz = -\int_{0}^{+\infty} g(z)dz + \int_{0}^{+\infty} g(z)dz = 0$$

$V(Z) = \int_{-\infty}^{+\infty} z^2\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)dz$ llamando: $u = z$ y $dv = z.\exp\left(-\frac{1}{2}z^2\right)$

resulta que: $du = dz$ y $v = -\exp\left(-\frac{1}{2}z^2\right)$. Por lo tanto, integrando por partes:

$$V(Z) = -\frac{1}{\sqrt{2\pi}}z.\exp\left(\frac{-1}{2}z^2\right)\Bigg|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\left(-\frac{1}{2}z^2\right)dz = 1$$

2.- Si $X \sim N(\mu,\sigma^2) \Rightarrow E(X) = \mu$ y $V(X) = \sigma^2$.

Sea $Z = \frac{X-\mu}{\sigma} \Rightarrow Z \sim N(0,1) \Rightarrow E(Z) = 0$ y $V(Z) = 1 \Rightarrow$

$E(X) = E(\sigma Z+\mu) = \sigma.E(Z)+\mu = \mu$

$V(X) = V(\sigma Z+\mu) = \sigma^2.V(Z) = \sigma^2$.

F. Distribución exponencial:

Se dice que ***X*** se distribuye como una exponencial de parámetro $\lambda > 0$ y se nota $X \sim \varepsilon(\lambda)$, si la función de densidad es:

$$f_X(x) = \lambda.e^{-\lambda x}.I_{[0,+\infty)}(x)$$

Veamos la forma de la distribución exponencial para los distintos valores del parámetro λ.

Proposición:

1.- Si $X \sim \varepsilon(\lambda) \Rightarrow E(X) = 1/\lambda$ y $V(X) = 1/\lambda^2$

Dem:

$$E(X) = \int_0^{+\infty} x\lambda\, e^{-\lambda x}dx = -xe^{-\lambda x}\Big|_0^{+\infty} - \int_0^{+\infty} -e^{-\lambda x}dx = \frac{-1}{\lambda}e^{-\lambda x}\Big|_0^{+\infty} = \frac{1}{\lambda}$$

Para la varianza debemos calcular:

$$E\left(X^2\right) = \int_0^{+\infty} x^2 \lambda\, e^{-\lambda x} dx = \frac{2}{\lambda^2} \text{ (integrando dos veces por partes)}$$

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

2.- Si $X \sim \varepsilon(\lambda) \Rightarrow F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$

Dem:

$$F_X(x) = \int_{-\infty}^{x} \lambda . e^{-\lambda t} . I_{[0,+\infty)}(t)\, dt = \int_0^x \lambda . e^{-\lambda t}\, dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x} \quad (\text{si } x \geq 0)$$

$$F_X(x) = \int_{-\infty}^{x} \lambda . e^{-\lambda t} . I_{[0,+\infty)}(t)\, dt = \int_{-\infty}^{x} \lambda . e^{-\lambda t} . 0\, dt = 0 \quad (\text{si } x < 0)$$

<u>Ejemplo 28:</u>

Suponga que el tiempo de respuesta ***X*** de una terminal conectada en línea tiene distribución exponencial con esperanza 5 seg.

a.- ¿Cuál es la probabilidad de que la respuesta tarde más de 10 segundos?

$$E(X) = 1/\lambda = 5 \Rightarrow \lambda = 1/5$$

$$P(X \geq 10) = 1 - F(10) = 1 - \left(1 - e^{-10/5}\right) = e^{-2} = 0.135$$

b.- ¿Cuál es la probabilidad de que la respuesta tarde entre 5 y 10 segundos?

$$P(5 \leq x \leq 10) = F(10) - F(5) = 1 - e^{-10/5} - \left(1 - e^{-5/5}\right) = e^{-1} - e^{-2} = 0.368 - 0.135 = 0.233$$

Propiedades de la función exponencial:

La distribución exponencial se usa para modelar el tiempo entre la ocurrencia de dos eventos sucesivos, tal como llamadas en un conmutador o llegada de clientes a un servicio. Esto se debe a que la distribución exponencial y la de Poisson están fuertemente relacionadas.

Propiedad: falta de memoria de la exponencial:

Si $T \sim \varepsilon(\lambda) \Rightarrow P\left(T \geq t + t_0 \,/\, T \geq t_0\right) = P(T \geq t)$

Nota:

Esta igualdad indica que no hay deterioro con el paso del tiempo, o sea que, la distribución del tiempo restante de vida *t* es independiente de la edad actual. Cuando en un sistema no se cumple esta propiedad no puede ser modelado con una exponencial y se requiere el uso de otras distribuciones más generales como la gamma, Weibull, o la distribución lognormal

Dem:

Sea $T \sim \varepsilon(\lambda) \Rightarrow$

$$P\left(T \geq t + t_0 \big/ T \geq t_0 \right) = \frac{P(T \geq t + t_0 \wedge T \geq t_0)}{P(T \geq t_0)} = \frac{P(T \geq t + t_0)}{P(T \geq t_0)} = \frac{1 - \left[1 - e^{-\lambda(t+t_0)}\right]}{1 - \left[1 - e^{-\lambda t_0}\right]} =$$

$$= \frac{e^{-\lambda(t+t_0)}}{e^{-\lambda t_0}} = e^{-\lambda t} = 1 - \left[1 - e^{-\lambda t}\right] = P(T \geq t)$$