Programa

1. **Introducción a las ecuaciones diferenciales (0.2 créditos)**. Generalidades, definición y clasificación. Conceptos de existencia, unicidad y tipos de soluciones.
2. **Ecuaciones diferenciales ordinarias (EDO) de primer orden (0.8 créditos)**. Definición y significado geométrico. Ecuaciones exactas, variables separadas. Factores integrantes. Ecuaciones separables y lineales. Métodos de transformación: ecuaciones homogéneas y de Bernoulli.
3. **EDO de orden superior (1 crédito)**. Reducción de orden. Ecuaciones lineales. Dependencia e independencia lineal de funciones. Ecuaciones lineales homogéneas: sistema fundamental de soluciones y fórmula de Liouville. Ecuaciones lineales completas: variación de constantes. Delta de Dirac: soluciones elementales y funciones generalizadas. Distribuciones.
4. **Sistemas de EDO (0.8 créditos)**. Reducción a una EDO. Integrales primeras. Sistemas lineales homogéneos y completos. Exponenciales de matrices.
5. **Transformación de Laplace (0.8 créditos)**. Definición y propiedades básicas. Convolución. Aplicación a la resolución de problemas de valores iniciales para EDO y sistemas de EDO con coeficientes constantes.
6. **Soluciones por series de potencias (1.2 créditos)**. Puntos ordinarios y singulares regulares. El método de Frobenius. Algunas funciones especiales: polinomios de Hermite y de Legendre, funciones de Bessel.
7. **Métodos Aproximados**.
8. **EDO no lineales y teoría de la estabilidad (1.2 créditos)**. Concepto de estabilidad. Puntos de equilibrio. Estabilidad lineal alrededor de los puntos de equilibrio. Espacio de fases.
9. **Teoría de Sturm-Liouville y funciones de Green (2 créditos)**. Espacios de funciones. Funciones ortogonales. Desarrollos en conjuntos de funciones ortogonales. Problemas de contorno. Teoría de Sturm-Liouville. Series de Fourier.
10. **Ecuaciones en derivadas parciales (EDP) (2 créditos)**. Introducción a las EDP. Problemas de contorno y separación de variables. Método de las Características de EDP.
11. **Probabilidad (1.2 créditos)**. Introducción a la probabilidad. Distribuciones discretas básicas. Distribuciones de probabilidad. Funciones de variable aleatoria. Teorema del límite central.
12. **Estadística (0.8 créditos)**. Concepto de estadístico. Estimadores. Estimación por intervalos de confianza.

Bibliografía

[1] J. M. Aguirregabiria, *Ecuaciones diferenciales ordinarias para estudiantes de Física*, UPV/EHU (2000).

[2] R. Haberman, *Ecuaciones en derivadas parciales con series de Fourier y problemas de contorno, 3.ªed.*, Prentice Hall (2003).

[3] A. I. Kiseliov, M. L. Krasnov y G. I. Makarenko, *Problemas de ecuaciones diferenciales ordinarias*, Mir-Rubiños 1860 (1992).

[4] K. F. Riley, M. P. Hobson y S. J. Bence, *Mathematical Methods for Physics and Engineering, 3rd ed.*, Cambridge (2003).

[5] W. E. Boyce y R. C. DiPrima, *Ecuaciones diferenciales y problemas con valores en la frontera, 4.ªed.*, Limusa (1998).

[6] F. Ayres, *Ecuaciones diferenciales*, Schaum McGraw-Hill (2001).

[7] G. B. Arfken y H. J. Weber, *Mathematical Methods for Physicists, 4th ed.*, Academic Press (1995).

[8] R. Courant y D. Hilbert, *Methods of Mathematical Physics (2 vols.)*, Wiley (1962).

[9] M. H. Holmes, *Introduction to Perturbation Methods*, Springer (1995).

[10] E. L. Ince, *Ordinary Differential Equations*, Dover (1956).

[11] L. M. Jones, *An introduction to Mathematical Methods of Physics*, Benjamin (1979).

[12] J. Mathews y R. L. Walker, *Mathematical Methods of Physics*, Benjamin (1970).

[13] W. Walter, *Ordinary Differential Equations*, Springer (1998).

[14] M. R. Spiegel y L. Abellanas, *Fórmulas y Tablas de Matemática Aplicada*, Schaum McGraw-Hill (1999).

[15] I. Bronshtein y K. Semendiaev, *Manual de Matemáticas*, Mir (1993).

[16] I. S. Gradshteyn y I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press (2014).

Aula 1.3

L	M	X	J
15-16	14-15	15-16	15-16

5 Abril
12 Abril
26 Abril
3 Mayo
} 12h-13h

Libro de Texto =>

Mathematical Methods for Physics and Engineering

Third Edition

K. F. RILEY, M. P. HOBSON and S. J. BENCE

Cambridge University Press

Probabilidad 1.2 créditos = 12 clases
Estadística 0.8 créditos = 8 clases

* Tutorías -> quedar conmigo.
* Ejercicios => plataforma eGela.

Por encima

Generating Functions = no

Binomial
Poisson
Gaussiana

-> 30.12 Joint distributions por encima (claro que existe)

resto del capítulo no se da

Statistics

31.1
31.2 } Bien

31.3 { Estimation of several quantities => no

31.4 31.4.1 Population mean
31.4.2 Population variance

CONCEPTOS BÁSICOS

- Un experimento aleatorio es aquél en el cual no disponemos de los conocimientos para predecir su resultado.
- El conjunto total de todos los posibles resultados de un experimento aleatorio se denomina **Espacio muestral** E
- Los conjuntos de resultados del experimento aleatorio que cumplen alguna propiedad se denominan **sucesos** o **eventos**.
- Ejemplos de experimentos aleatorios y su espacio muestral correspondiente

Experimentos aleatorios $\Rightarrow$ Lánzese un dado de seis caras $\Rightarrow E = \{1, 2, 3, 4, 5, 6\}$

Lánzese una moneda $\Rightarrow E = \{+, -\}$

¿Cuál es el coche más rápido del mundo?

Colóquese un radar en una carretera y mídase la velocidad de los coches $\Rightarrow E = [0 \text{ km/h}, 350 \text{ km/h}]$

En los dos primeros ejemplos E es un conjunto finito mientras que en el último ejemplo E es infinito.

- Ejemplos de sucesos:
  - para el experimento aleatorio del dado: $A \equiv \{\text{El resultado es múltiplo de } 2\} = \{2, 4, 6\}$

    $B \equiv \{\text{El resultado es divisible por } 3\} = \{3, 6\}$
  - para el experimento aleatorio del radar en la carretera: $A \equiv \{\text{Coches multados en una autopista española}\} = \left(120 \frac{\text{km}}{\text{h}}, 350 \frac{\text{km}}{\text{h}}\right)$

En el primer ejemplo A y B son conjuntos finitos mientras que en el segundo ejemplo A es un conjunto infinito. Nótese que el propio espacio muestral E también define un suceso.

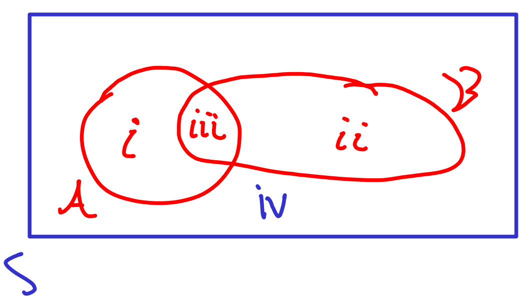
* Diagramas de Venn. El espacio muestral y los sucesos de un experimento aleatorio se pueden representar utilizando una representación gráfica conocida como "diagrama de Venn"

Ejemplo: experimento del dado: $S=\{1,2,3,4,5,6\}$. Definimos los sucesos:

$A=\{$El resultado es múltiplo de 2$\}=\{2,4,6\}$

$B\equiv\{$El resultado es divisible por 3$\}=\{3,6\}$

Diagrama de Venn = Rectángulo que representa el espacio muestral. Dentro se dibujan una o más curvas cerradas cuyos interiores representan los sucesos.



Formulamos ahora las siguientes definiciones:

i ≡ resultados que pertenecen al suceso A pero no al suceso B.

ii ≡ Resultados que pertenecen al suceso B pero no al suceso A.

iii ≡ Resultados que pertenecen a ambos sucesos A y B.

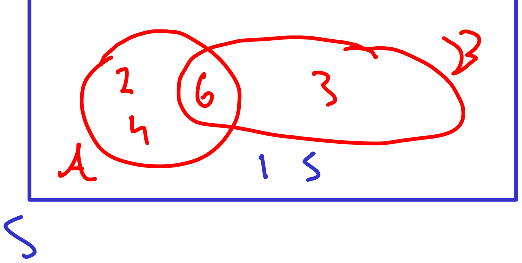
iv ≡ Resultados que no pertenecen ni al suceso A ni al suceso B.

En el ejemplo se tiene que:

i $=\{2,4\}$, ii $=\{3\}$,

iii $=\{6\}$, iv $=\{1,5\}$

Por tanto en este caso se tiene el siguiente diagrama de Venn:



$A\cup B=\{2,3,4,6\}$

$A\cap B=\{6\}$

Nótese que los conjuntos de resultados definidos en i)-iv) son ellos mismos sucesos. Algunos de estos sucesos reciben nombres especiales:

* Operaciones con sucesos:

* Se dice que los resultados que cumplen iii) pertenecen a la **intersección** de los sucesos A y B, denotada por $A \cap B$

* Se dice que los resultados que cumplen i)–iii) pertenecen a la **unión** de los sucesos A y B, que se denota por $A \cup B$.

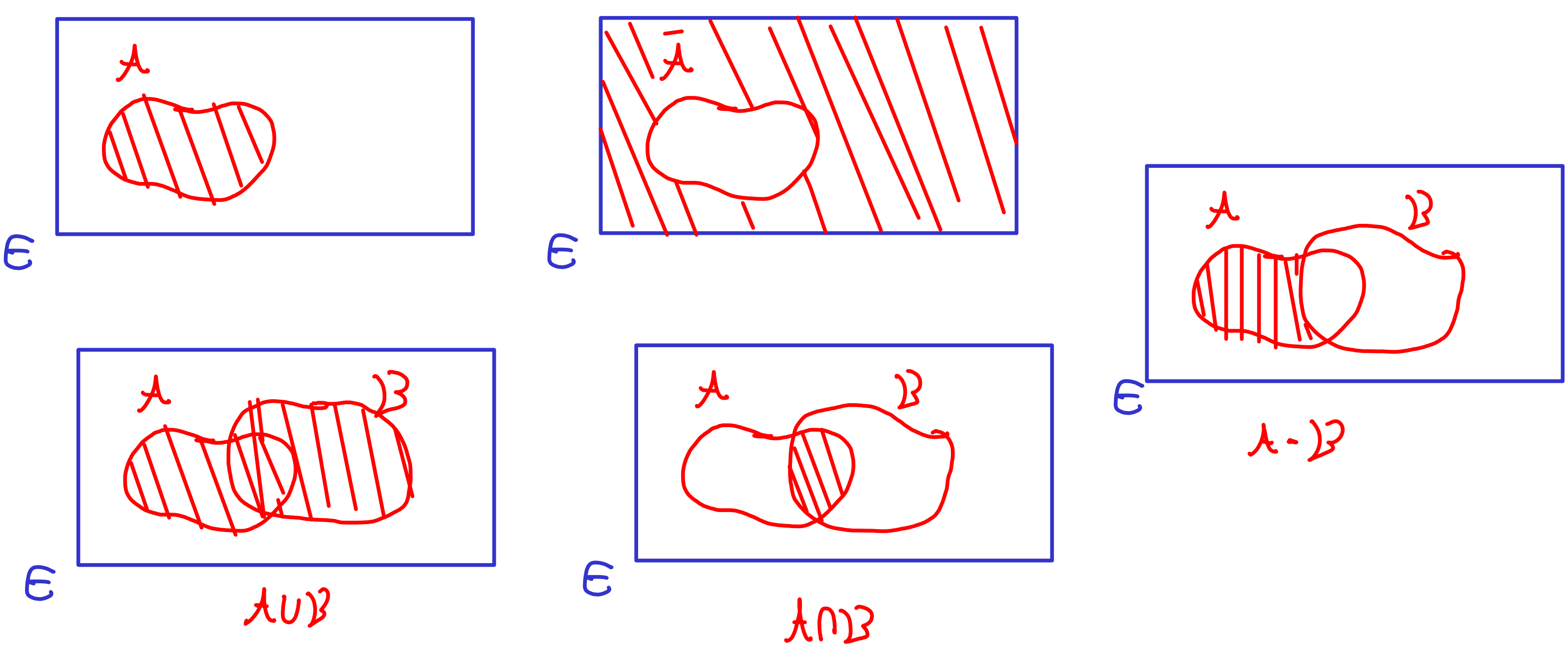
Si ningún resultado cumple iii) entonces se dice que los sucesos A y B son **excluyentes o disjuntos**.

* Un suceso que no contiene ningún resultado se denomina **suceso vacío**, denotado por ϕ. Por tanto A y B son excluyentes si $A \cap B = \phi$

* Los resultados que no pertenecen a un suceso dado A definen un nuevo suceso denotado por $\bar{A}$ (complementario de A). Claramente $A \cap \bar{A} = \phi$ y $A \cup \bar{A} = S$

* Los resultados que pertenecen al suceso A pero no al suceso B definen un nuevo suceso denotado por $A - B$ (a veces la notación es A/B).
Se tiene la propiedad $A - B = A \cap \bar{B}$

Representación de las operaciones con sucesos mediante diagramas de Venn



* Operaciones con sucesos: álgebra de sucesos

Utilizando diagramas de Venn o estudiando ejemplos particulares es posible ver que la unión, la intersección de sucesos y el complementario de un suceso cumplen las siguientes propiedades: (Recordamos primero las definiciones:)

- $E \equiv \{$ Conjunto de todos los resultados de un experimento aleatorio (espacio muestral) $\}$
- $\phi \equiv \{$ Conjunto vacío (suceso que no contiene ningún resultado) $\}$
- A, B, C, conjuntos de resultados de un experimento aleatorio (sucesos).

- Existencia de identidades para cada operación: $A \cup \phi = A \ , \ A \cap E = A$

- Ambas operaciones son asociativas:
$$A \cup (B \cup C) = (A \cup B) \cup C$$
$$A \cap (B \cap C) = (A \cap B) \cap C$$

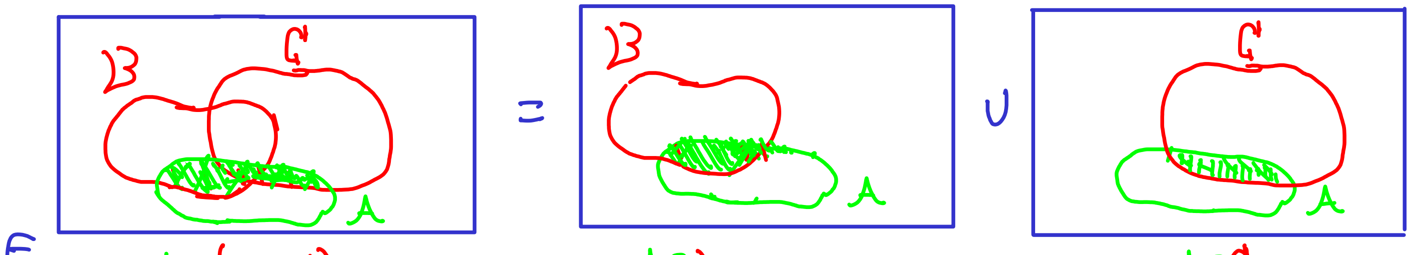
- Ambas operaciones son conmutativas: $A \cup B = B \cup A \ , \ A \cap B = B \cap A$

- Cada operación es distributiva respecto de la otra:
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

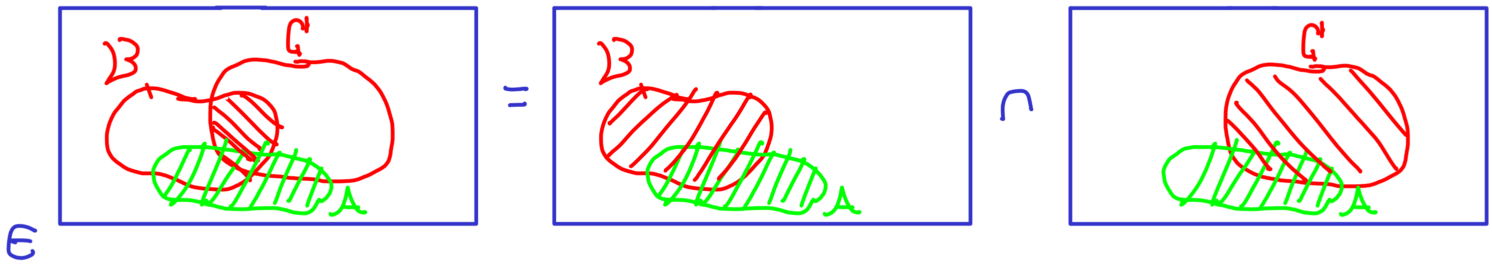
- Para todo suceso A, existe su complementario $\bar{A}$ caracterizado por $A \cup \bar{A} = E \ , \ A \cap \bar{A} = \phi$

El complementario es único.

Ejemplos de propiedades sencillas:

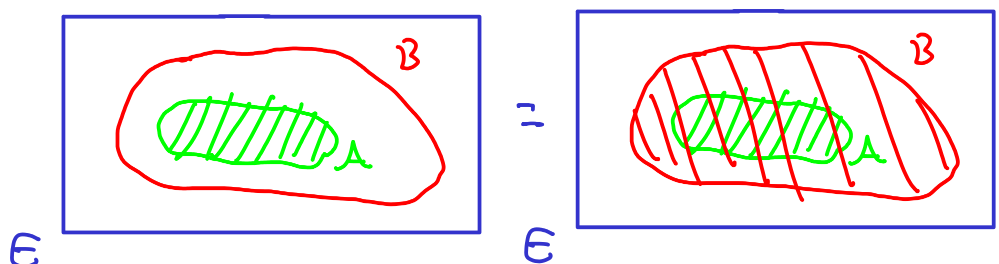


$A \cap (B \cup C) = A \cap B \cup A \cap C$



$A \cup (B \cap C) = A \cup B \cap A \cup C$

$A = A \cap B \Leftrightarrow B = A \cup B$

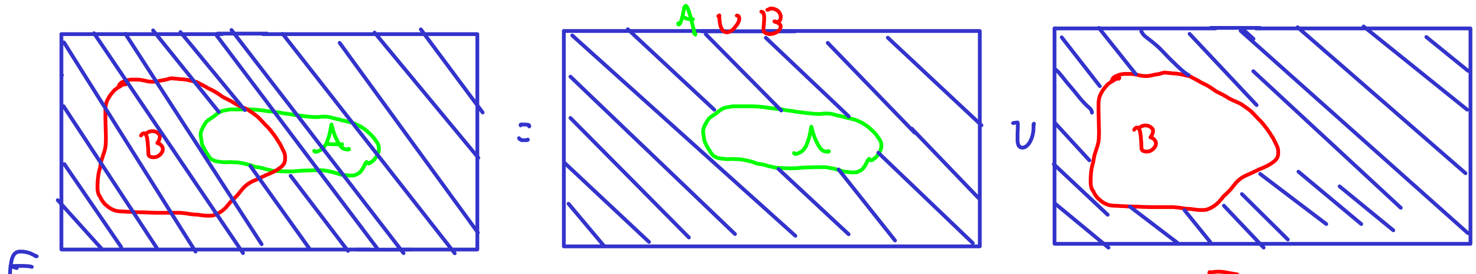


$A = A \cap B$ $=$ $A \cup B = B$

Leyes de de Morgan

$\overline{A \cup B}$ $=$ $\bar{A}$ $\cap$ $\bar{B}$

$\overline{A \cup B}$



$\overline{A \cap B}$ $=$ $\bar{A}$ $\cup$ $\bar{B}$

* Operación de inclusión de sucesos.

* Si todos los resultados de un suceso A también pertenecen al suceso B entonces escribiremos $A \subset B$. Puede haber resultados en B que no pertenezcan al suceso A. La operación "$\subset$" se puede definir en términos de $\cap$ así: $A \subset B \Leftrightarrow A = A \cap B$

Se tiene que $A \subset B$ y $B \subset A$ si y sólo si $A = B$

► *Show that* (i) $\underbrace{A \cup (A \cap B) = A \cap (A \cup B) = A}_{\text{Las leyes de absorción}}$, (ii) $(A - B) \cup (A \cap B) = A$.

i)

$$A \cup (A \cap B) = (A \cap E) \cup (A \cap B) = A \cap (E \cup B) = A \cap E = A$$

utilizar

propiedad distributiva

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \rightarrow (A \cap B) \cup (A \cap C) = A \cap (B \cup C)$$

$A = A$, $B = E$, $C = B$, $A = A$, $E = E$, $B = B$

$$A \cap (A \cup B) = (A \cup A) \cup (A \cap B) = A \cup (A \cap B) = A$$

Ley distributiva

usamos la propiedad que acabamos de probar.

ii) $$(A - B) \cup (A \cap B) = (A \cap \bar{B}) \cup (A \cap B) = A \cap (\bar{B} \cup B) = A \cap E = A$$

sacamos A factor común

► *There exist two events A and B such that*

$$\overline{(X \cup A)} \cup \overline{(X \cup \bar{A})} = B.$$

Find an expression for the event X in terms of A and B.

$\Rightarrow$ De Morgan $(X \cup A) \cap (X \cup \bar{A}) = \bar{B}$

Sacar X factor común $\quad X \cup (A \cap \bar{A}) = \bar{B}$

$$X \cup \phi = \bar{B} \Rightarrow X = \bar{B}$$

MATERIAL SUPLEMENTARIO (FUERA DEL PROGRAMA)

ÁLGEBRAS DE BOOLE

Las propiedades explicadas de los sucesos y el espacio muestral en relación a las operaciones "$\cup$" y "$\cap$" constituyen un caso particular de una construcción abstracta de la teoría de conjuntos conocida como álgebra de Boole

Definición formal de álgebra de Boole: Un álgebra de Boole $\mathcal{B}$ es un 6-conjunto $\mathcal{B} = (M, e, \phi, \cup, \cap, \bar{\ })$ donde M es un conjunto $\cup$, $\cap$ son operaciones binarias definidas en M, $\bar{\ }$ es una operación unaria definida en M y $e \in M$, $\phi \in M$ son elementos de M que cumplen los siguientes axiomas:

- ϕ y e son elementos identidad para $\cup$ y $\cap$ respectivamente: $a \cup \phi = a$, $a \cap e = a$ $\forall a \in M$
- $\cup$ y $\cap$ son operaciones asociativas: $(a \cup b) \cup c = a \cup (b \cup c)$, $\forall a, b, c \in M$. $(a \cap b) \cap c = a \cap (b \cap c)$, $\forall a, b, c \in M$.
- ambas operaciones son conmutativas: $a \cup b = b \cup a$, $a \cap b = b \cap a$ $\forall a, b \in M$.
- cada operación es distributiva respecto de la otra: $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$, $\forall a, b, c \in M$. $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$, $\forall a, b, c \in M$.
- $\forall a \in M$ $\exists$ $\bar{a} \in M$ tal que $a \cup \bar{a} = e$, $a \cap \bar{a} = \phi$

Proposición $\forall a\in M$ se tiene $a\cup e = a$, $a\cap \phi = \phi$

Prueba:

$$a\cup e \overset{\text{axiomas}}{=} (a\cup e)\cap e \overset{\text{axiomas}}{=} (a\cup e)\cap(a\cup\bar a) \overset{\text{propiedad distributiva}}{=} a\cup(e\cap\bar a) \overset{\text{axiomas}}{=} a\cup\bar a \overset{\text{axiomas}}{=} e$$

$$a\cap\phi \overset{\text{axiomas}}{=} (a\cap\phi)\cup\phi \overset{\text{axiomas}}{=} (a\cap\phi)\cup(a\cap\bar a) \overset{\text{propiedad distributiva}}{=} a\cap(\phi\cup\bar a) \overset{\text{axiomas}}{=} a\cap\bar a \overset{\text{axiomas}}{=} \phi$$

En particular si en la proposición anterior tomamos $a = e, \phi$ se obtiene

$$\left.\begin{array}{ll} e\cup e = e & \phi\cap\phi = \phi \\ \phi\cup e = e & e\cap\phi = \phi \end{array}\right\} \Rightarrow$$

combinando estos resultados con la unicidad del complementario probada más abajo $\Rightarrow \bar e = \phi$

Proposición (leyes de absorción) $\forall a,b\in M$ $\quad a\cup(a\cap b) = a$, $a\cap(a\cup b) = a$

Prueba:

$$a\cup(a\cap b) = (a\cap e)\cup(a\cap b) = a\cap(e\cup b) = a\cap e = a$$

utilizar

$$(\underset{a}{A}\cap \underset{e}{B})\cup(\underset{b}{A}\cap \underset{a}{C}) = \underset{e}{A}\cap(\underset{b}{B}\cup C)$$

$$a\cap(a\cup b) \overset{\text{Ley distributiva}}{=} (a\cap a)\cup(a\cap b) = a\cup(a\cap b) = a$$

usamos la propiedad que acabamos de probar.

Proposición (leyes de la tautología) $\cup$ y $\cap$ son operaciones idempotentes: $\forall a,b\in M$ se tiene que $a\cup a = a$, $a\cap a = a$ $\forall a\in M$

Prueba: Tómese $b = e, \phi$ en las leyes de absorción y aplíquese la proposición anterior:

$$a\cup(a\cap e) = a \Rightarrow a\cup a = a$$

$$a\cap(a\cup\phi) = a \Rightarrow a\cap a = a$$

Proposición

$$a = a \cap b \text{ si y sólo si } a \cup b = b$$

Prueba: $\Rightarrow$) supongamos que $a = a\cap b$. Entonces $a \cup b = (a \cap b) \cup b \underset{\text{propiedad distributiva}}{=} (a\cup b) \cap (b \cup b) \underset{\text{tautología}}{=}$
$= (a\cup b) \cap b = b \cap (a \cup b) = b \cap (b \cup a) = b$
ley de absorción

$\Leftarrow$) Supongamos que $a \cup b = b$. Entonces $a \cap b = a \cap (a \cup b) \underset{\text{propiedad distributiva}}{=} (a \cap a) \cup (a \cap b) \underset{\text{tautología}}{=}$
$a \cup (a \cap b) = a$
ley de absorción

Utilizando este resultado podemos definir una operación binaria $\subseteq$ del siguiente modo:

$$a \subseteq b \iff a = a \cap b$$

Proposición: La operación $\subseteq$ es un orden parcial en el conjunto M

Prueba: para ver que $\subseteq$ es un orden parcial tenemos que comprobar lo siguiente:

i) La operación $\subseteq$ es reflexiva. Por la tautología $a = a \cap a \Rightarrow a \subseteq a$

ii) La operación $\subseteq$ es antisimétrica: hay que probar que si $a \subseteq b$ y $b \subseteq a$ entonces $a = b$ $\forall a, b \in M$

$$\left.\begin{array}{l} a \subseteq b \Leftrightarrow a = a \cap b \\ b \subseteq a \Leftrightarrow b = b \cap a \end{array}\right\} \Rightarrow a = b \quad (\cap \text{ es conmutativa})$$

iii) La operación $\subseteq$ es transitiva: hay que probar que si $a \subseteq b$ y $b \subseteq d$ entonces $a \subseteq d$, $\forall a, b, d \in M$

$$\left.\begin{array}{l} a \subseteq b \Leftrightarrow a = a \cap b \\ b \subseteq d \Leftrightarrow b = b \cap d \end{array}\right\} \begin{array}{l} a = a \cap (b \cap d) = (a \cap b) \cap d = a \cap d \Rightarrow \\ \Rightarrow a \subseteq d \end{array}$$

Proposición $\forall a \in M$ se tiene que $\phi \leq a \leq e$

Prueba $a = a \cap e$ (uno de los axiomas) $\Rightarrow a \leq e$

$a = a \cup \phi$ (uno de los axiomas) $\Rightarrow \phi = \phi \cap a$ (usar la proposición) $\Rightarrow$

$\Rightarrow \phi \leq a$.

Lema (unicidad del complementario) si $\left\{\begin{array}{l} A \cup B = A \cup C = e \\ A \cap B = A \cap C = \phi \end{array}\right\} \Rightarrow$ entonces $B = C = \bar{A}$

Prueba

$$\left.\begin{array}{l} B = B \cap e = B \cap (A \cup C) = (B \cap A) \cup (B \cap C) = \phi \cup (B \cap C) = B \cap C \\ C = C \cap e = C \cap (A \cup B) = (C \cap A) \cup (C \cap B) = \phi \cup (C \cap B) = C \cap B \end{array}\right\} \Rightarrow B = C = \bar{A}$$

Proposición (Leyes de De Morgan)

$$\left.\begin{array}{l} i)\ \overline{a \cup b} = \bar{a} \cap \bar{b} \\ ii)\ \overline{a \cap b} = \bar{a} \cup \bar{b} \end{array}\right\}$$

propiedad distributiva

$$\left.\begin{array}{l} i)\ (a \cup b) \cap (\bar{a} \cap \bar{b}) = (a \cap \bar{a} \cap \bar{b}) \cup (b \cap \bar{a} \cap \bar{b}) = \phi \cup \phi = \phi \\ (a \cup b) \cup (\bar{a} \cap \bar{b}) = (a \cup b \cup \bar{a}) \cap (a \cup b \cup \bar{b}) = e \cap e = e \end{array}\right\}$$

utilizar la unicidad del complementario para concluir que $\overline{a \cup b} = \bar{a} \cap \bar{b}$

ii)

$$\left.\begin{array}{l} (\bar{a} \cup \bar{b}) \cap (a \cap b) = (\bar{a} \cap a \cap b) \cup (\bar{b} \cap a \cap b) = \phi \cup \phi = \phi \\ (\bar{a} \cup \bar{b}) \cup (a \cap b) = (\bar{a} \cup \bar{b} \cup a) \cap (\bar{a} \cup \bar{b} \cup b) = e \cap e = e \end{array}\right\} \Rightarrow (\overline{a \cap b}) = \bar{a} \cup \bar{b}$$

propiedad distributiva

(unicidad del complementario)

Proposición $a \leq b \Rightarrow \bar{b} \leq \bar{a}$

Prueba Si $a \leq b$ entonces $a \cap b = a \Rightarrow a \cup b = b$ (absorción) $\Rightarrow \bar{a} \cap \bar{b} = \bar{b}$ (de Morgan)

$\Rightarrow \bar{b} \leq \bar{a}$

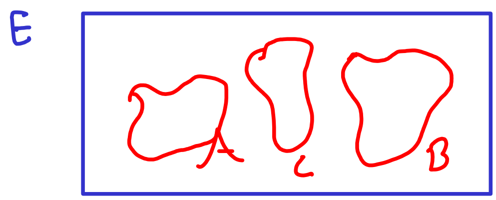
Definición Dados dos elementos $a, b \in M$ definimos su diferencia así:

$$a - b = a \cap \overline{b}$$ (elementos contenidos en "a" y no en "b")

Claramente $e - a = e \cap \overline{a} = \overline{a}$, $a - e = a \cap \overline{e} = a \cap \phi = \phi$

PROBABILIDAD

De manera más formal: tenemos un espacio muestral E y su álgebra (booleana) de sucesos asociada



$\Rightarrow$ $M \equiv$ conjunto de sucesos que podemos definir a partir de E. ¿Cuál es el tamaño de un suceso?

$\Rightarrow$ A cada suceso del álgebra le asignaremos un número positivo entre 0 y 1 que denominaremos probabilidad de dicho suceso. Este número mide "el tamaño" del suceso

$\Rightarrow$ Definimos una aplicación $p: M \to [0,1]$ $\forall A \in M$ $p(A)$ es la probabilidad del suceso A.

* La aplicación p cumple los siguientes axiomas (axiomas de Kolmogorov)

i) $p(A) \geq 0 \quad \forall A \in M$

ii) $p(E) = 1$

Si tenemos varios sucesos disjuntos $A, B, C,$, entonces

iii) $p(A \cup B \cup C) = p(A) + p(B) + p(C)$

$\Rightarrow$ Podemos pensar que p es una especie de área (medida) que definimos para cada suceso con relación a E

* Consecuencias básicas de los axiomas (se pueden ver dibujando diagramas de Venn o de manera formal).

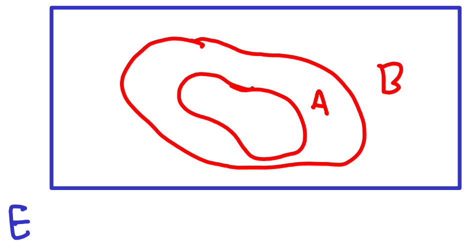
i) $p(\emptyset) = 0 \Rightarrow$ Si A es un suceso cualquiera, A y $\emptyset$ son disjuntos $A \cap \emptyset = \emptyset \Rightarrow$

$p(A \cup \emptyset) = p(A) = p(A) + p(\emptyset) \Rightarrow p(\emptyset) = 0$

En un experimento aleatorio siempre obtendremos un resultado por lo que la probabilidad de obtener un suceso que no tiene ningún resultado es 0

c) $A \subset B \Rightarrow p(A) \le p(B)$ Prueba: A , $B\setminus A$ son disjuntos $(A\cap(B\setminus A) = A\cap B\cap \bar{A} = \emptyset)$

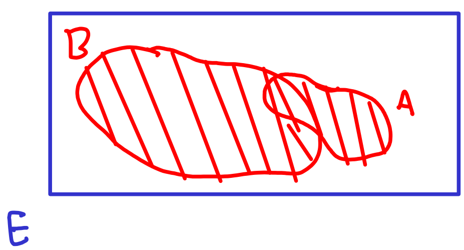
y $B = A \cup B\setminus A \Rightarrow p(B) = p(A) + p(B\setminus A) \Rightarrow p(B) \ge p(A)$



$\Rightarrow$ el area de B es mayor o igual que el area de A.

iii) Si A y B no son disjuntos entonces $p(A\cup B) = \underbrace{p(A) + p(B) - p(A\cap B)}$

sumar las areas y restar el area de la intersección.



Prueba formal:

Definimos los sucesos disjuntos $H_1 \equiv A$, $H_2 \equiv B \setminus (A\cap B) = B\setminus A$.

$B\setminus(A\cap B) = B\cap(\overline{A\cap B}) = B\cap(\bar{A}\cup\bar{B}) = (B\cap\bar{A})\cup(\underbrace{B\cup\bar{B}}_{E}) = B\cap\bar{A} = B\setminus A$

$H_1\cap H_2 = A\cap B\cap\bar{A} = \emptyset$ y $H_1\cup H_2 = A\cup(B\cap\bar{A}) = (A\cup B)\cap E = A\cup B$

$\Rightarrow$ usar los axiomas $p(H_1\cup H_2) = p(A\cup B) = p(A) + p(B\setminus(A\cap B))$

Definimos los sucesos disjuntos $Z_1 \equiv A\cap B$, $Z_2 \equiv B\setminus(A\cap B) \Rightarrow Z_1\cap Z_2 = \emptyset$, $Z_1\cup Z_2 = B$

$\Rightarrow$ usar los axiomas $p(Z_1\cup Z_2) = p(B) = p(A\cap B) + p(B\setminus(A\cap B))$

Eliminar $p(B\setminus A\cap B)$ de las dos ecuaciones anteriores $\Rightarrow p(A\cup B) - p(B) = p(A) - p(A\cap B)$

Supongamos que tenemos un experimento aleatorio determinado que repetimos N veces.

- La frecuencia absoluta de un suceso S se define como el número de veces n_s que obtenemos un resultado perteneciente a dicho suceso.

Ejemplo: tiramos una moneda al aire 18 veces y obtenemos cara 8 veces. Para el suceso $S=$"obtener cara" la frecuencia absoluta es $n_s = 8$.

- La frecuencia relativa de un suceso S se define como el número de veces n_s que obtenemos un resultado perteneciente a dicho suceso dividido entre el número de veces N que hemos hecho el experimento: $f_r(S) \equiv \frac{n_s}{N}$

Ejemplo: colocamos un radar en una autopista y medimos la velocidad de los coches que pasan. Considere el suceso:

$S \equiv$ "La velocidad del coche está en el intervalo $\left[100 \frac{km}{h}, 130 \frac{km}{h}\right]$"

Repetimos el experimento aleatorio 100 veces y encontramos 83 resultados que están en el suceso S que acabamos de definir.

$$n_s = \text{frecuencia absoluta de } S = 83$$

$$f_r(S) = \text{Frecuencia relativa de } S = \frac{n_s}{N} = \frac{83}{100}$$

- La idea intuitiva es que si el número de veces N que repetimos un experimento aleatorio tiende a infinito, entonces la frecuencia relativa de un suceso S debe estabilizarse en torno a la "probabilidad del suceso S".

Experimento aleatorio = "tirar una moneda al aire y ver el resultado" $S = \{\text{obtener cara}\}$

$N \to \infty$ Frecuencia relativa de $S \sim \frac{1}{2}$

En general:

$$f_r(S) \xrightarrow[N\to\infty]{} p(S) \equiv \text{Probabilidad del suceso } S$$

Nótese que no tenemos un límite en el sentido matemático de la palabra. Se trata de una relación de "proximidad" que se estudiará y cuantificará en la parte de estadística a través de un proceso de "inferencia estadística".

- La frecuencia relativa cumple las siguientes propiedades que coinciden con las propiedades de la probabilidad:

i) $0 \le f_r(S) \le 1$, $\forall$ suceso S (recuérdese la definición $f_r(S) = 0 \le \frac{n_S}{N} \le 1$)

ii) $f_r(E) = 1$ $\left(f_r(E) = \frac{N_E}{N} = \frac{N}{N} = 1 \right)$

iii) $f_r(\phi) = 0$ $\left(f_r(\phi) = \frac{N_\phi}{N} = \frac{0}{N} = 0 \right)$

iv) Si A y B son dos sucesos cualesquiera entonces

$$f_r(A\cup B) = \frac{n_{A\cup B}}{N} = \frac{n_A + n_B - n_{A\cap B}}{N} = \frac{n_A}{N} + \frac{n_B}{N} - \frac{n_{A\cap B}}{N} =$$

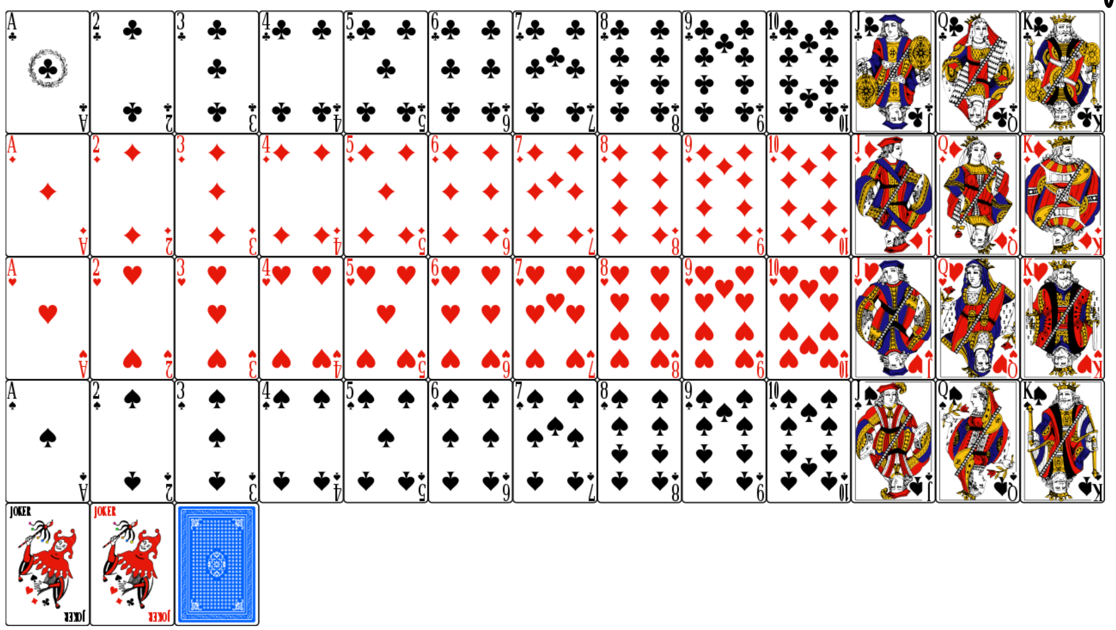
$= f_r(A) + f_r(B) - f_r(A\cap B)$. Esto implica que si $A\cap B = \phi$ entonces $f_r(A\cup B) = f_r(A) + f_r(B)$.

FRENCH DECK OF CARDS

4 suits (palos) = clubs, diamonds, hearts and spades

Ace   Jack Queen King

52 cards



Clubs (tréboles)

diamonds (diamantes)

Hearts (corazones)

Spades (picas)

Jokers
Comodines

* Compute the probability of getting an ace or a jack when drawing a card at random from a deck of cards (52 cards, no jokers)

Random experiment: draw a card from a deck of cards

$A = \{\text{we get an ace}\}$
$B = \{\text{we get a jack}\}$ $\Rightarrow p(A \cup B) = p(A) + p(B)$

$A \cap B = \emptyset \Rightarrow$ excluding events

Decompose A and B into elementary and exclusive events (= outcomes)

$A = \{\text{ace of clubs}\} \cup \{\text{ace of diamonds}\} \cup \{\text{ace of hearts}\} \cup \{\text{ace of spades}\}$

$B = \{\text{jack of clubs}\} \cup \{\text{jack of diamonds}\} \cup \{\text{jack of hearts}\} \cup \{\text{jack of spades}\}$

$p(\text{elementary event}) = \dfrac{1}{52} \Rightarrow p(A) = \dfrac{4}{52} = \dfrac{2}{26} = \dfrac{1}{13} = p(B) \Rightarrow p(A \cup B) = \dfrac{2}{13}$

* Ejemplos: Considérese el experimento aleatorio de lanzar una moneda

$E = \{+, -\}$ $+$: "sale cara" => Suceso; $-$: "sale cruz" => suceso } sucesos excluyentes

$E = \{+\} \cup \{-\} \Rightarrow p(E) = 1 = p(+) + p(-)$

Si el dado no está trucado entonces todos los resultados deben tener la misma probabilidad. =>

$p(+) = p(-)$

$\Rightarrow p(+) = p(-) = \frac{1}{2}$

* Calcúlese la probabilidad de extraer un as o una sota al extraer una carta de una baraja española.

Experimento aleatorio = extracción de una carta de la baraja española

$A = \{$el resultado es un as$\}$

$B = \{$el resultado es una sota$\}$

=> Queremos calcular $p(A \cup B) = p(A) + p(B)$

$A \cap B = \emptyset \Rightarrow$ sucesos excluyentes

Descomponemos A y B en sucesos elementales (= resultados) y excluyentes.

$A = \{$as de oros$\} \cup \{$as de copas$\} \cup \{$as de bastos$\} \cup \{$as de espadas$\}$

$B = \{$sota de oros$\} \cup \{$sota de copas$\} \cup \{$sota de bastos$\} \cup \{$sota de espadas$\}$

Como la baraja española tiene 40 cartas, si la baraja no está trucada entonces la probabilidad de cada suceso elemental es $\frac{1}{40}$ y por tanto

$p(\{$as de oros$\}) = p(\{$as de copas$\}) = p(\{$as de bastos$\}) = p(\{$as de espadas$\}) = \frac{1}{40}$

$p(\{$sota de oros$\}) = p(\{$sota de copas$\}) = p(\{$sota de bastos$\}) = p(\{$sota de espadas$\}) = \frac{1}{40}$

Así que:

$$\left.\begin{array}{l} p(A) = \dfrac{4}{40} = \dfrac{1}{10} \\[2ex] p(B) = \dfrac{4}{40} = \dfrac{1}{10} \end{array}\right\} \quad p(A \cup B) = p(A) + p(B) = \frac{2}{10} = \frac{1}{5}$$

* Calcular la probabilidad de que al extraer una carta de una baraja se obtenga un as o que la carta sea oros.

$A = \{\text{la carta es un as}\}$

$B = \{\text{la carta es oros}\}$

$A \cap B = \{\text{la carta es un as y es oros}\} = \{\text{la carta es el as de oros}\}$

Descomponemos en sucesos elementales:

$$B = \{\text{la carta es oros}\} = \{\text{la carta es as de oros}\} \cup \{\text{la carta es 2 de oros}\} \cup \{\text{la carta es 3 de oros}\} \cup \dots$$

$$\dots \cup \{\text{la carta es rey de oros}\} \Rightarrow p(A) = p\{\text{la carta es oros}\} = \frac{10}{40} = \frac{1}{4}$$

$$p(A) = p\{\text{la carta es un as}\} = \frac{4}{40} = \frac{1}{10} \quad \text{(ejemplo anterior)}$$

$$p(A \cap B) = p\{\text{la carta es el as de oros}\} = \frac{1}{40} \quad \text{(suceso elemental)}$$

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = \frac{1}{4} + \frac{1}{10} - \frac{1}{40} = \frac{10+4-1}{40} = \frac{13}{40}$$

* Las probabilidades respectivas de obtener 1, 2, 3, 4, 5, 6 al lanzar un dado trucado son $\frac{1}{2}k, k, k, k, k, 2k$. Calcúlese el valor de k.

Los sucesos $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$ y $\{6\}$ son sucesos excluyentes y su unión coincide con el espacio muestral. Así que:

$$1 = p(\{1\}) + p(\{2\}) + p(\{3\}) + p(\{4\}) + p(\{5\}) + p(\{6\}) = \frac{1}{2}k + k + k + k + k + 2k = \frac{1}{2}k + 6k =$$

$$= \frac{13}{2}k \Rightarrow k = \frac{2}{13}$$

* Si hay más de dos sucesos asociados a un experimento aleatorio y no son excluyentes entonces debemos proceder de manera inductiva:

$$p(A_1 \cup \underbrace{A_2 \cup A_3}_{B}) = p(A_1 \cup B) = p(A_1) + p(B) - p(A_1 \cap B)$$

$$= p(A_1) + p(B) - p(A_1 \cap (A_2 \cup A_3))$$

$$= p(A_1) + p(B) - p((A_1 \cap A_2) \cup (A_1 \cap A_3)) =$$

$$= p(A_1) + p(A_2 \cup A_3) - p(A_1 \cap A_2) - p(A_1 \cap A_3) + p(A_1 \cap A_2 \cap A_1 \cap A_3)$$

$$= p(A_1) + p(A_2) + p(A_3) - p(A_2 \cap A_3) - p(A_1 \cap A_2) - p(A_1 \cap A_3) + p(A_1 \cap A_2 \cap A_3)$$

La fórmula anterior se generaliza de la manera siguiente:

$$p(A_1 \cup \dots \cup A_n) = \sum_{i=1}^{n} p(A_i) - \sum_{i<j}^{n} p(A_i \cap A_j) + \sum_{i<j<k} p(A_i \cap A_j \cap A_k) \dots$$

$$+ \dots + (-1)^{n+1} p(A_1 \cap \dots \cap A_n)$$

Prueba: usar inducción en n.

* Calcúlese la probabilidad de que al sacar una carta de una baraja española se obtenga un as, la carta sea copas o la carta presente un número impar (1, 3, 5, 7)

$A_1 = \{\text{sale un as}\}$

$A_2 = \{\text{sale copas}\}$

$A_3 = \{\text{la carta presenta un número impar}\}$

$A_1 \cap A_2 = \{\text{sale un as y sale copas}\} = \{\text{la carta es el as de copas}\}$

$A_1 \cap A_3 = \left\{\begin{array}{l}\text{sale un as y la} \\ \text{carta presenta un} \\ \text{número impar}\end{array}\right\} = \{\text{la carta es un as}\} = A_1$

$A_2 \cap A_3 = \left\{\begin{array}{l}\text{sale copas y la} \\ \text{carta presenta un} \\ \text{número impar}\end{array}\right\} = \left\{\begin{array}{l}\text{as de copas} \\ \text{3 de copas} \\ \text{5 de copas} \\ \text{7 de copas}\end{array}\right\}$, $A_1 \cap A_2 \cap A_3 = \left\{\begin{array}{l}\text{sale un as y sale copas} \\ \text{y la carta presenta un} \\ \text{número impar}\end{array}\right\} = \left\{\begin{array}{l}\text{as de} \\ \text{copas}\end{array}\right\}$

$p(A_1) = \frac{4}{40} = \frac{1}{10} = p(A_1 \cap A_3) = \frac{1 \text{ as} \times 4 \text{ palos}}{40 \text{ cartas}} = \frac{4 \text{ ases}}{40 \text{ cartas}} = \frac{1}{10}$

$\qquad\qquad\qquad\qquad\qquad \parallel$

$p(A_2) = \frac{10}{40} = \frac{1}{4} \qquad p(A_2 \cap A_3) = \frac{\text{as copas} + 3 \text{ copas} + 5 \text{ copas} + 7 \text{ copas}}{40 \text{ cartas}} = \frac{4}{40} = \frac{1}{10}$

$p(A_3) = \frac{4 \text{ numeraciones impares} \times 4 \text{ palos}}{40 \text{ cartas}} = \frac{4 \cdot 4}{40} = \frac{16}{40} = \frac{2}{5}$

$p(A_1 \cap A_2) = \frac{1}{40} = p(A_1 \cap A_2 \cap A_3)$,

Aplicamos la fórmula:

$$p(A_1 \cup A_2 \cup A_3) = \frac{1}{10} + \frac{1}{4} + \frac{2}{5} - \cancel{\frac{1}{40}} - \frac{1}{10} - \frac{1}{10} + \cancel{\frac{1}{40}} = \frac{2+5+8-2-2}{20} = \frac{11}{20}$$

MATERIAL SUPLEMENTARIO (FUERA DEL PROGRAMA)

- Axiomas de Kolmogorov. Sea $\mathcal{B}$ un álgebra booleana de sucesos asociada a un experimento aleatorio y denotemos por $M \subset \mathcal{B}$ el espacio de sucesos (no confundir el espacio de sucesos M con el espacio muestral E que es el espacio de resultados). Para todo suceso $A \in M$ definimos un número real $p(A)$ de acuerdo con los siguientes axiomas:

Pregunta: ¿ $\mathcal{P}(E) = M$ siempre?

i) $p(A) \geq 0, \ \forall A \in M$

ii) $p(E) = 1$

iii) Si $\{A_i\}_{i \in J}$ es una familia numerable de sucesos disjuntos $A_i \cap A_j = \emptyset, \ \forall i, j \in J$ entonces

$$p\left(\bigcup_{i \in J} A_i\right) = \sum_{i \in J} p(A_i).$$

Claramente p es una aplicación de la forma $p: M \to \mathbb{R}^+$. La terna (E, M, p) se denomina espacio de probabilidad. De los axiomas se deduce el siguiente resultado:

Proposición La aplicación $p: M \to \mathbb{R}^+$ cumple las siguientes propiedades

i) $p(\emptyset) = 0$

ii) Si $A, B \in M$ son tales que $A \subseteq B$ entonces $p(A) \leq p(B)$

iii) $p(A) \leq 1, \ \forall A \in M$

Prueba Para un suceso A definimos la familia numerable $\{H_i\}_{i \in J}$ donde $H_1 = A$, $H_2 = B \setminus A$, $H_i = \emptyset \ \forall i \geq 3$, siendo B tal que $A \subseteq B$ y $p(B) \neq 0$.

$B \setminus A = B \cap \bar{A}$ así que $(B \setminus A) \cap A = \phi$ y además $H_i \cap \phi = \phi$

Se tiene entonces que $\bigcup_{i \in J} H_i = A \cup (B \setminus A) = A \cup (B \cap \bar{A}) = (A \cup B) \cap (A \cup \bar{A}) =$ Lo usar que $A \subseteq B$

$= B \cap E = B$. Por tanto, según los axiomas de Kolmogorov:

$$p(B) = p\left(\bigcup_{i \in J} H_i\right) = \sum_{i \in J} p(H_i) = p(A) + p(B \setminus A) + \sum_{i=3}^{\infty} p(\phi)$$

La serie $\sum_{i=3}^{\infty} p(\phi)$ tiene que converger a un número positivo finito y como $p(\phi) \geq 0$ la única posibilidad es que $p(\phi) = 0$. Además utilizando de nuevo la positividad deducimos que $p(B) \geq p(A)$ lo que prueba el punto (i). Finalmente, dado que para todo suceso A se tiene que $A \subseteq E$ deducimos que $p(A) \leq p(E) = 1$.

Proposición Para dos sucesos cualesquiera A, B se tiene que

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Prueba Definimos la familia de sucesos $H_1 = A$, $H_2 = B \setminus (A \cap B)$

$B \setminus (A \cap B) = B \cap \overline{(A \cap B)} = B \cap (\bar{A} \cup \bar{B}) = (B \cap \bar{A}) \cup (E) = B \cap \bar{A} = B \setminus A$

así que $H_1 \cap H_2 = A \cap B \cap \bar{A} = \phi$ y $H_1 \cup H_2 = A \cup (B \cap \bar{A}) = (A \cup B) \cap E = A \cup B$

Por tanto, según los axiomas

(a) $p(H_1 \cup H_2) = p(A \cup B) = p(A) + p(B \setminus (A \cap B))$

Si ahora definimos la familia $Z_1 = A \cap B$, $Z_2 = B \setminus (A \cap B)$ entonces $Z_1 \cap Z_2 = \phi$ y $Z_1 \cup Z_2 = B$ así que

(b) $p(Z_1 \cup Z_2) = p(B) = p(A \cap B) + p(B \setminus (A \cap B))$

Restando (a) y (b) obtenemos

$$p(A\cup B)-p(B)=p(A)-p(A\cap B) \Rightarrow p(A\cup B)=p(A)+p(B)-p(A\cap B)$$

Un espacio de probabilidad es un caso particular de una estructura más general conocida como "Espacio de medida". Para definir un espacio de medida necesitamos primero la noción de "Σ-álgebra". Una Σ-álgebra es un conjunto M de elementos que es "cerrado" para las operaciones $\cup$ (unión) y $^{-}$ (complementario). Es decir:

$\forall a,b\in M \quad a\cup b\in M$

$\forall a\in M,\ \bar{a}\in M.$

Un álgebra de Boole (y en particular el álgebra de sucesos asociada a un experimento aleatorio) son ejemplos de Σ-álgebras.

Definición

* Una medida sobre una Σ-álgebra M es una aplicación $\mu: M\to \mathbb{R}^+$ que cumple los siguientes axiomas.

i) $\mu(\phi)=0$

ii) si $\{a_i\}_{i\in J}$ es una familia numerable de elementos de Σ tal que $a_i\cap a_j=\phi \quad \forall i,j\in J$, entonces se tiene que

$$\mu\Big(\bigcup_{i\in J} a_i\Big)=\sum_{i\in J}\mu(a_i)$$

El par (M,μ) se denomina "espacio de medida"

* Un caso particular de espacio de medida es un espacio de probabilidad (E,M,p)

* Si el espacio muestral es finito entonces se puede descomponer en la unión disjunta de sus elementos, que son sucesos mutuamente excluyentes:

$$E = \bigcup_i \{s_i\} \qquad \{s_i\} = \text{sucesos elementales} \quad \{s_i\} \cap \{s_j\} = \phi$$

Los sucesos elementales también se denominan <u>casos posibles</u>

En este caso, cualquier suceso A será la unión disjunta de algunos de los sucesos elementales, denominándose dichos sucesos <u>casos favorables</u> al suceso A

$$A = \bigcup_j \{A_j\}$$

* En un experimento aleatorio ideal podemos hacer la suposición de que los sucesos elementales son equiprobables, es decir

$$p(s_i) = p(s_j) \quad \forall i, j \quad \Rightarrow$$ basta saber la probabilidad de uno de los sucesos elementales, llamémosle s_0

Por tanto $\quad 1 = \sum_i p(s_i) = N\, p(s_0) \qquad N \equiv$ número de casos posibles.

$$\Rightarrow p(s_0) = \frac{1}{N}$$

De aquí deducimos que para un suceso arbitrario A

$$p(A) = \sum_j p(A_j) = \sum_j p(s_0) = m\, p(s_0) = \frac{m}{N} \qquad m = \text{número de casos favorables}$$

* Llegamos así a la clásica fórmula de la probabilidad de un suceso asociado a un experimento aleatorio cuyo espacio muestral tiene un número finito de elementos:

Fórmula de Laplace

$$\boxed{p(A) = \frac{\text{número de casos favorables a } A}{\text{número de casos posibles}}}$$

Veamos algunos de los ejemplos anteriores:

* Calcúlese la probabilidad de extraer un as o una sota al extraer una carta de una baraja española.

$$\left.\begin{array}{l}\text{número de casos favorables} = 4 \text{ ases} + 4 \text{ sotas} = 8 \\ \text{número de casos posibles} = 40 \text{ cartas}\end{array}\right\} p = \frac{8}{40} = \frac{1}{5}$$

* Calcúlese la probabilidad de que al sacar una carta de una baraja española se obtenga un as, la carta sea copas o la carta presente un número impar (1, 3, 5, 7)

$$\text{número de casos favorables} = \left\{\begin{array}{l} 4 \text{ ases} \\ \quad + \\ 10 \text{ copas} - 1 \text{ as de copas} \\ \\ \left(\begin{array}{l}4 \text{ numeraciones} \\ \text{impares}\end{array} - 1 \text{ as}\right) \times (4 \text{ palos} - 1 \text{ palo de copas}) \end{array}\right.$$

$= 4 + 9 + 3 \times 3 = 22$

número de casos posibles $= 40$ cartas $\qquad p = \frac{22}{40} = \frac{11}{20}$

Probabilidad condicionada

* Existen situaciones en las que tenemos más de un experimento aleatorio. Normalmente los experimentos aleatorios se efectúan de manera consecutiva y el resultado de uno de ellos puede estar influenciado (condicionado) por los experimentos anteriores.

Ejemplo

Se extraen 2 cartas al azar de una baraja española: ¿Cuál es la probabilidad de extraer 2 ases si:

i) La primera carta extraída se pone aparte y después se extrae la segunda carta 2

ii) La primera carta extraída se devuelve a la baraja y después se extrae la 2ª carta.

- Imaginemos 2 experimentos aleatorios

- Experimento aleatorio X_1 con espacio muestral E_{X_1}

- Definimos

$A_{X_1} \equiv$ suceso asociado al experimento aleatorio X_1 con espacio muestral E_{X_1}

- Experimento aleatorio X_2 con espacio muestral E_{X_2} relacionado (o no) con E_{X_1}

$B_{X_2} \equiv$ suceso asociado al experimento aleatorio X_2 con espacio muestral E_{X_2}

Por definición $\quad p(A_{X_1} \cap B_{X_2}) = p(A_{X_1}) \cdot p(B_{X_2})$

- En muchos casos B_{X_2} está relacionado con A_{X_1}. Cuando eso pasa se dice que B_{X_2} está condicionado por el suceso A_{X_1}. En ese caso, en vez de escribir $p(B_{X_2})$ escribimos $p(B_{X_2}/A_{X_1})$ (Creo que esto debería ser $p_{A_{X_1}}(B_{X_2})$)

- En general no se hace referencia a los experimentos aleatorios X_1 y X_2 en las fórmulas. Se escriben simplemente los sucesos sin subíndices.

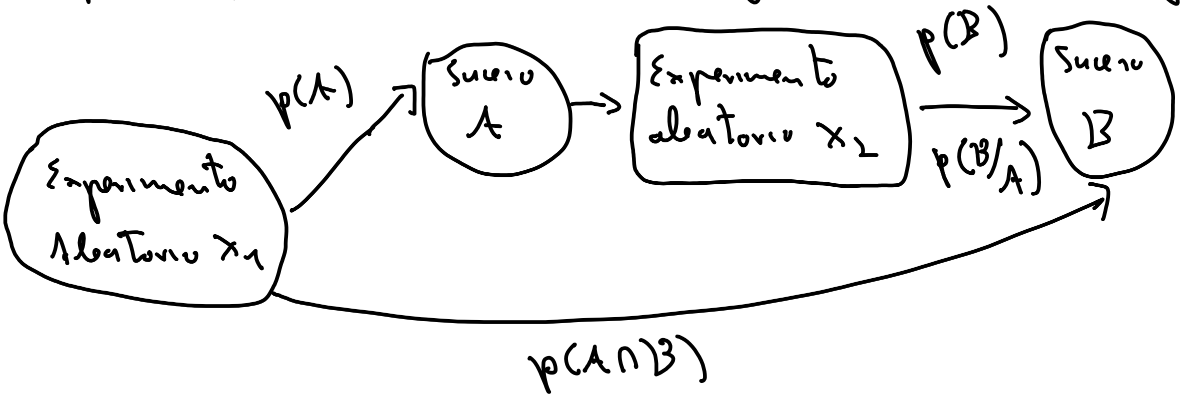
$p(B/A) \equiv$ probabilidad del suceso B condicionado por el suceso A

$$\boxed{p(A \cap B) = p(A) \cdot p(B/A) \iff p(B/A) = \frac{p(A \cap B)}{p(A)}}$$

- Vemos que la fórmula se reduce al caso ya estudiado si sólo tenemos un experimento aleatorio

Si $A = E \Rightarrow \quad p(B/E) = \frac{p(B \cap E)}{p(E)} = p(B)$

La fórmula $p(A\cap B)=p(A)\cdot p(B/A)$ admite la siguiente "representación gráfica"



* La probabilidad de $A\cap B$ es la probabilidad de recorrer el camino que va desde el principio del diagrama hasta el final y se obtiene multiplicando las probabilidades de cada uno de los tramos

* Si los experimentos aleatorios X_1 y X_2 son independientes entonces podemos considerar que los sucesos correspondientes son independientes y no están relacionados. En ese caso $p(B/A)=p(B)$, $p(A/B)=p(A)$ y se tiene que

$$p(A\cap B)=p(A)\cdot p(B)$$

* Dos sucesos A y B son estadísticamente independientes si

$$p(A\cap B)=p(A)\,p(B)$$

Los sucesos A y B pueden estar asociados a dos experimentos aleatorios que pueden ser iguales o diferentes.

<u>Definición</u> Dos sucesos $\{A_1,\dots A_n\}$ asociados a n experimentos aleatorios (iguales o diferentes) son estadísticamente independientes si se tiene que

$$p(A_1\cap\dots\cap A_n)=p(A_1)\cdots p(A_n)$$

MATERIAL SUPLEMENTARIO (FUERA DEL PROGRAMA)

¿Qué hemos hecho desde el punto de vista formal?

Hemos partido de un espacio de probabilidad (E, M, p) y definimos un nuevo "espacio de probabilidad reducido" (A, M_A, p_A) donde $A \subset E$, $A \in M$ y $M_A \subset M$. Entonces $\forall x \in M_A$ tenemos que

$$p_A(x) \equiv \frac{p(x \cap A)}{p(A)}$$

Pregunta: ¿$p_A(\bar{x}) = 1 - p_A(x)$? Respuesta: no tiene que ser el "complementario respecto a A" no respecto al espacio muestral completo.

Ejercicio: Demostrar que con la definición anterior (A, M_A, p_A) es realmente un espacio de probabilidad (se cumplen los axiomas de Kolmogorov)

Solución

primer axioma: $p_A(x) = \dfrac{p(x \cap A)}{p(A)} \geq 0 \quad \forall x \in M_A$

segundo axioma: $p_A(A) = \dfrac{p(A \cap A)}{p(A)} = \dfrac{p(A)}{p(A)} = 1$

tercer axioma: $\{x_i\}_{i \in J}$ familia numerable de sucesos disjuntos tales que $x_i \cap x_j = \phi \quad \forall i, j$

$$p_A\left(\bigcup_{i \in J} x_i\right) = \frac{p\left[\left(\bigcup_{i \in J} x_i\right) \cap A\right]}{p(A)} = \frac{p\left(\bigcup_{i \in J} (x_i \cap A)\right)}{p(A)} = \sum_{i \in J} \frac{p(x_i \cap A)}{p(A)} =$$

$$= \sum_{i \in J} p_A(x_i)$$

usar $(x_i \cap A) \cap (x_j \cap A) = \phi$

Analizamos ahora el ejemplo planteado anteriormente

Se extraen 2 cartas al azar de una baraja española: ¿Cuál es la probabilidad de extraer 2 ases si:

i) La primera carta extraída se pone aparte y después se extrae la segunda carta 2

ii) La primera carta extraída se devuelve a la baraja y después se extrae la 2ª carta.

Tenemos dos experimentos aleatorios consecutivos $\Rightarrow$ $\begin{cases} 1^{\circ} \text{ extracción de una carta de la baraja} \Rightarrow X_1 \\ 2^{\circ} \text{ extracción de una segunda carta del montón resultante} \Rightarrow X_2 \end{cases}$

Sucesos:

$A = \{$la carta es un as en la 1ª extracción$\}$ $B = \{$la carta es un as en la 2ª extracción$\}$

i) Si la 1ª carta no se devuelve los experimentos aleatorios están relacionados y por tanto debemos calcular $p(B/A)$ y $p(A)$ usando la fórmula de Laplace:

$$p(A) = \frac{\text{nº de ases}}{\text{nº total de cartas}} = \frac{4}{40} = \frac{1}{10}. \qquad \underset{\substack{\uparrow \\ \text{usar Laplace}}}{p(B/A) =} \frac{\text{nº de ases tras quitar un as}}{\text{nº de cartas tras quitar una.}} = \frac{3}{39} = \frac{1}{13}$$

Por tanto $\Rightarrow$ probabilidad de sacar 2 ases $= p(A \cap B) = p(A)\, p(B/A) = \frac{1}{10} \cdot \frac{1}{13} = \frac{1}{130}$

ii) Si se devuelve la primera carta extraída entonces los experimentos aleatorios son independientes $\Rightarrow p(B) = p(B/A)$

$\Rightarrow$ los sucesos A y B son estadísticamente independientes

$$p(A \cap B) = p(A)\, p(B/A) = p(A) \cdot p(B) = \frac{1}{10} \times \frac{1}{10} = \frac{1}{100}$$

$$p(A) = \frac{\text{nº de ases}}{\text{nº total de cartas}} = \frac{4}{40} = \frac{1}{10} \qquad p(B) = \frac{\text{nº de ases}}{\text{nº total de cartas}} = \frac{4}{40} = \frac{1}{10}$$

El problema se puede también resolver por conteo directo

Podemos juntar los 2 experimentos aleatorios sucesivos y considerar que son un solo experimento.

Experimento aleatorio = extracción de dos cartas de una baraja española.

Nos importa el orden: {as de bastos, sota de copas} ≠ {sota de copas, as de bastos}

$A \equiv$ {La primera carta extraída de la baraja es un as}

$B \equiv$ {La segunda carta extraída de la baraja es un as}

$A \cap B =$ {La primera carta extraída de la baraja es un as y la segunda carta extraída de la baraja es un as}

i) $p(A \cap B) = \dfrac{4 \text{ ases} \times 3 \text{ ases descontando un as}}{\text{nº total de pares de cartas (importa el orden)}} = \dfrac{12}{40 \times 39} = \dfrac{6}{20 \cdot 39} = \dfrac{3}{390} = \dfrac{1}{130}$

ii) $p(A \cap B) = \dfrac{4 \text{ ases} \times 4 \text{ ases}}{\text{nº total de pares de cartas (importa el orden)}} = \dfrac{16}{40 \times 40} = \dfrac{1}{10^2} = \dfrac{1}{100}$

• Se saca una bola al azar de un bombo cuyas bolas están numeradas del 1 al 60 y a continuación se saca otra bola. ¿Cuál es la probabilidad de que la segunda bola tenga impreso el número 25 si la primera era una bola impar?

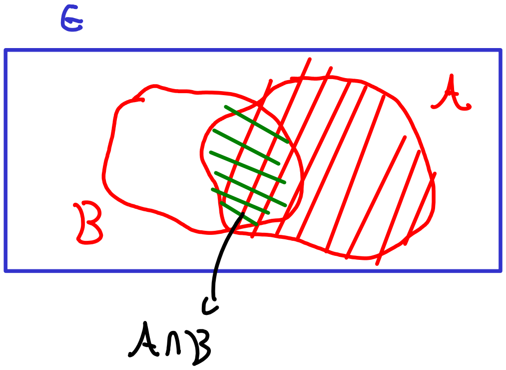
$A =$ {Sale una bola impar}

$B =$ {Sale la bola con el nº 25}

$p(B/A) = \dfrac{\text{nº de bolas con el 25}}{\text{nº de bolas impares que quedan en el bombo}} = \dfrac{1}{59}$

PENSAR SOBRE ESTO

Podemos considerar que $p(B/A)$ es una "probabilidad reducida" del suceso B en un espacio muestral reducido porque sabemos que A ya ha ocurrido.



Ejemplo: Poner como ejercicio.

Hallar la probabilidad de sacar dos ases en dos extracciones consecutivas de cartas de una baraja francesa si

i) La carta obtenida en la primera extracción se devuelve a la baraja antes de hacer la 2ª extracción.

ii) La carta obtenida en la primera extracción se pone aparte y después se hace la 2ª extracción

Caso i)

Experimento aleatorio = extracción consecutiva de 2 cartas de una baraja francesa.

nos importa el orden {as tréboles, 7 corazones} ≠ {7 corazones, as tréboles}

A = {la primera carta extraída es un as}

B = {la segunda carta extraída es un as}

$A \cap B = \{$ la primera carta extraída es un as y la segunda carta extraída es un as $\}$

Se tiene que $p(B/A) = \dfrac{p(A\cap B)}{p(A)} \Rightarrow p(A\cap B) = p(A)\,p(B/A)$

como hemos devuelto la carta sacada en la 1ª extracción a la baraja tenemos que

$p(B/A) = p(B)$ ya que A y B son independientes $\Rightarrow p(A\cap B) = p(A)\,p(B)$

$$p(A) = \frac{\text{nº de ases}}{\text{nº de cartas}} = \frac{4}{52} = p(B) \Rightarrow p(A\cap B) = \left(\frac{4}{52}\right)^2 = \frac{1}{13^2} = \frac{1}{169}$$

ii) El experimento aleatorio y los sucesos son los mismos pero ahora como ponemos aparte la carta de la primera extracción $p(B/A) \neq p(B)$

$$p(A) = \frac{\text{nº de ases}}{\text{nº de cartas}} = \frac{4}{52}$$

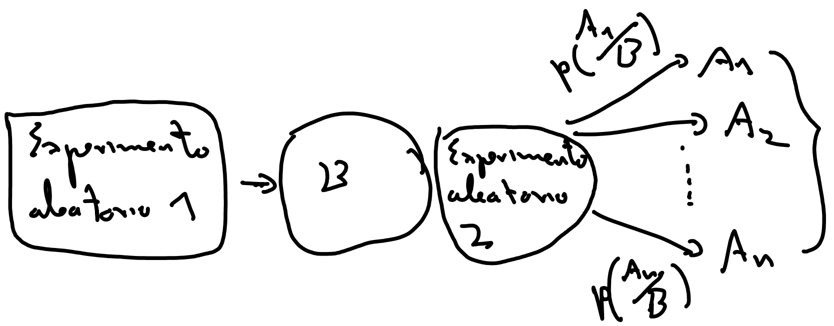
$$p(B/A) = \frac{\text{nº de ases si antes hemos sacado un as}}{\text{nº de cartas tras sacar una}} = \frac{3}{51}$$

$$p(A\cap B) = p(A)\cdot p(B/A) = \frac{4}{52}\cdot\frac{3}{51} = \frac{1}{13}\cdot\frac{3}{51} = \frac{3}{663} = \frac{1}{221}$$

* Si el suceso A es la unión de n sucesos A_i mútuamente excluyentes entonces para otro suceso cualquiera B se tiene que $\{B \cap A_i\}_{i=1}^{n}$ es también una familia de sucesos mutuamente excluyentes. Por tanto:

$$A = \bigcup_i A_i \Rightarrow p(A \cap B) = p\left[\bigcup_i (A_i \cap B)\right] = \sum_i p(A_i \cap B) \Rightarrow$$

$$\frac{p(A \cap B)}{p(B)} = p(A/B) = \sum_i p\frac{(A_i \cap B)}{p(B)} = \sum_i p(A_i/B)$$ (fórmula de adición de las probabilidades condicionales)



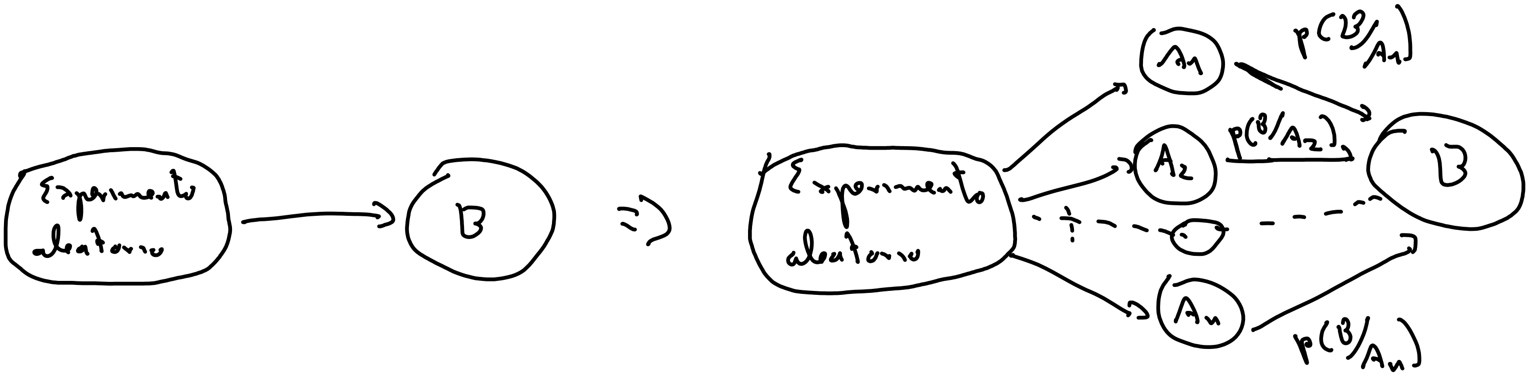
$\Rightarrow p(A/B) = \sum_{i=1}^{n} p(A_i/B)$

(sumar las probabilidades de cada una de las ramas)

* Si la unión de un conjunto de sucesos mutuamente excluyentes A_i es todo el espacio muestral entonces para cualquier suceso B se tiene que

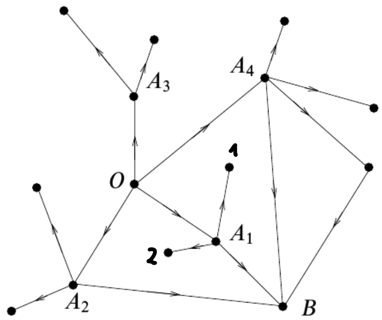
$$p(B) = p(B \cap \bigcup_i A_i) = p(\bigcup_i B \cap A_i) = \sum_i p(B \cap A_i) = \sum_i p(A_i)\, p(B/A_i)$$

(fórmula de completitud de la probabilidad condicionada)



Descomponemos el experimento aleatorio que da lugar al suceso B en dos: uno que da lugar a los sucesos $\{A_1, \ldots A_n\}$ y otro que da lugar al suceso B a partir de $A_1 \ldots A_n$. La probabilidad de llegar a B es la suma de las probabilidades de cada uno de los caminos que nos llevan a B

Ejemplo: Sea un conjunto de varias islas conectadas por puentes que sólo se pueden transitar en una dirección (véase la figura). Un conductor que se halla en una isla elige al azar la dirección que va a tomar para pasar a la siguiente isla. Si inicialmente se halla en O ¿cuál será la probabilidad de que el conductor llegue a B?



Experimento aleatorio: soltamos al conductor en O y dejamos que vaya pasando de una isla a otra hasta llegar a su destino

Este experimento aleatorio se descompone en sub-experimentos aleatorios que se definen así:

subexperimento aleatorio: el conductor que está en una isla escoge un puente para cruzar a la siguiente.

P. ej. el conductor que está en la isla O toma un puente para cruzar a otra isla:

$A_1 = \{$El puente va a la isla $A_1\}$ $\quad A_1 \cup A_2 \cup A_3 \cup A_4 = E_O$, $A_i \cap A_j = \phi$

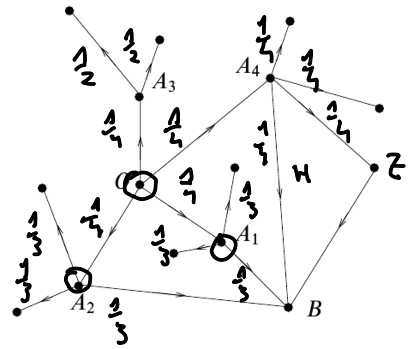
$A_2 = \{$El puente va a la isla $A_2\}$ $\quad$ Como los sucesos A_1, A_2, A_3, A_4 son

$A_3 = \{$El puente va a la isla $A_3\}$ $\quad$ equiprobables $\Rightarrow p(A_1) = p(A_2) = p(A_3) = p(A_4) =$

$A_4 = \{$El puente va a la isla $A_4\}$ $\quad = \frac{1}{4}$

El conductor que está en la isla A_1 toma un puente para cruzar a otra isla. Como a A_1 sólo se llega a través de O podemos considerar que este experimento aleatorio está condicionado por el anterior. Así podemos escribir los sucesos:

Así se deduce que cada nodo del diagrama da lugar a un experimento aleatorio y podemos asociar una probabilidad a cada uno de los segmentos que unen los nodos.



Para hallar la probabilidad que nos piden aplicamos la fórmula de la descomposición de la probabilidad condicionada:

$$p(B) = p(B/A_2)\, p(A_2) + p(B/A_1)\, p(A_1) + p(B/A_4)\, p(A_4) + p(B/A_3) \cdot p(A_3)$$

miramos al diagrama para completar los valores:

$$p(B/A_2) = \frac{1}{3} \qquad p(B/A_1) = \frac{1}{3} \qquad p(B/A_3) = 0$$

$$p(B/A_4) = p(B/Z) \cdot p(Z/A_4) + p(B/H) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$p(A_1) = \frac{1}{4} \qquad p(A_2) = \frac{1}{4} \qquad p(A_3) = \frac{1}{4} \qquad p(A_4) = \frac{1}{4}$$

$$p(B) = \frac{1}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{12} + \frac{1}{12} + \frac{1}{8} = \frac{1}{6} + \frac{1}{8} = \frac{4+3}{24} = \frac{7}{24}$$

* Teorema de Bayes

Si tenemos dos sucesos A y B entonces la fórmula de la probabilidad condicionada nos dice que:

$$p(A/B) = \frac{p(A \cap B)}{p(B)} \qquad p(B/A) = \frac{p(A \cap B)}{p(A)}$$

Podemos eliminar $p(A \cap B)$ de estas 2 ecuaciones:

$$\boxed{p(A/B)\,p(B) = p(B/A)\,p(A)} \quad \left(\text{teorema de Bayes}\right)$$

De aquí se deduce que $p(A/B) = p(B/A) \Rightarrow p(B) = p(A)$

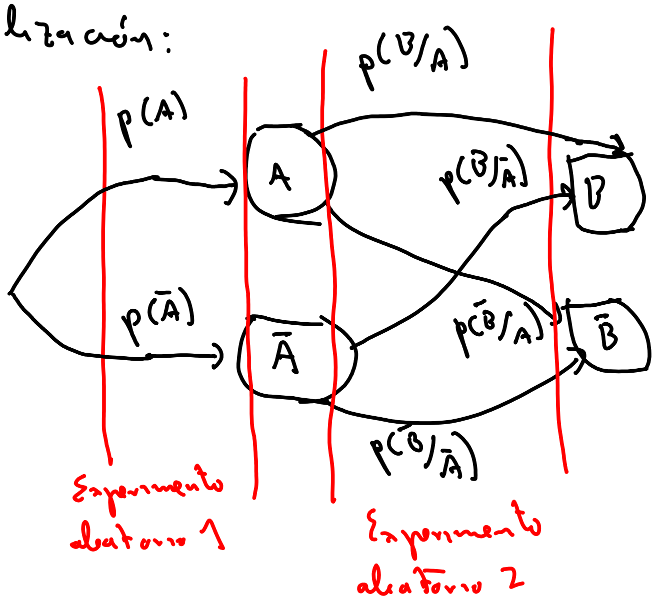
Si no conocemos $p(B)$ lo que se puede hacer es tomar $E = A \cup \bar{A}$ y usar la fórmula de la completitud de la probabilidad condicionada.

$$p(B) = p(A)\cdot p\left(\frac{B}{A}\right) + p(\bar{A})\cdot p\left(\frac{B}{\bar{A}}\right)$$ de donde

$$\boxed{p(A/B) = \frac{p(A)\cdot p(B/A)}{p(A)\cdot p\left(\frac{B}{A}\right) + p(\bar{A})\cdot p\left(\frac{B}{\bar{A}}\right)}} \quad \left(\text{Formulación alternativa del teorema de Bayes}\right)$$

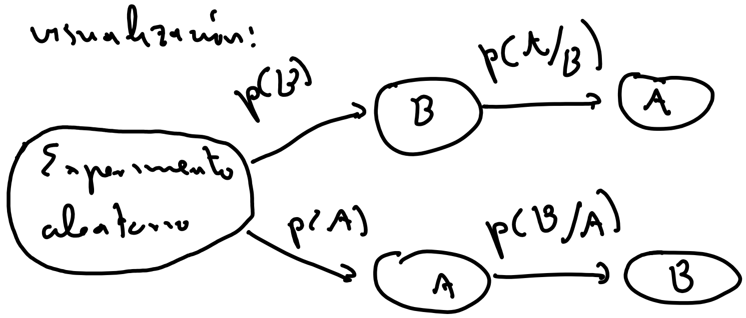
Esta fórmula admite una representación visual que se muestra en la página siguiente:

Visualización:



$p(A/B)$ es la probabilidad del camino que pasa por A para ir a B dividido entre la probabilidad de todos los caminos que pasan por A

Otra visualización:



Si hacemos 2 experimentos aleatorios la probabilidad de que salga A si primero salió B es la misma que la de que salga B si antes salió A

Supongamos que un análisis de sangre para una determinada enfermedad posee los siguientes índices de fiabilidad:

- Para las personas infectadas el test da positivo en el 99.99% de los casos.
- Para las personas no infectadas el test da positivo en el 0.02% de los casos.

Además, se sabe que una de cada 10.000 personas está infectada por la enfermedad. Se selecciona una persona de la población al azar, se le practica el test y da positivo ¿Cuál es la probabilidad de que esté infectada?

Respuesta: Experimentos aleatorios ⇒ Escógase una persona al azar y hágasele el test.

Sucesos:

A ≡ {la persona está infectada}   La probabilidad pedida es $p(A/B)$

B ≡ {el test da positivo}

Usamos la fórmula de Bayes para calcular la probabilidad

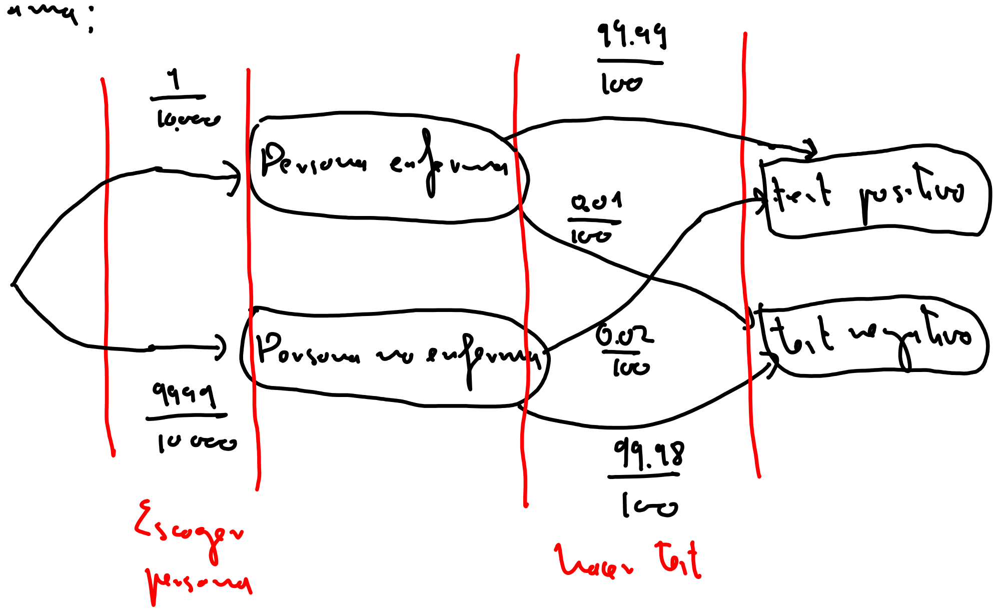
$$p(A/B) = \frac{p(A)\cdot p(B/A)}{p(A)\cdot p(B/A) + p(\bar{A})\cdot p(B/\bar{A})}$$

$p(A) = \frac{1}{10.000}$, $p(B/A) = \frac{99.99}{100}$, $p(B/\bar{A}) = \frac{0.02}{100}$ $\Rightarrow$

$$p(A/B) = \frac{\frac{1}{10.000}\cdot\frac{99.99}{100}}{\frac{1}{10.000}\cdot\frac{99.99}{100} + \left(\frac{9999}{10.000}\right)\frac{0.02}{100}} = \frac{99.99}{99.99 + 9999\cdot 0.02} = \frac{99.99}{299.97} \approx \frac{1}{3}$$

$$\begin{array}{r} 9999 \\ 0.02 \\ \hline 199.98 \\ 99.99 \\ \hline 299.97 \end{array}$$

Diagrama:

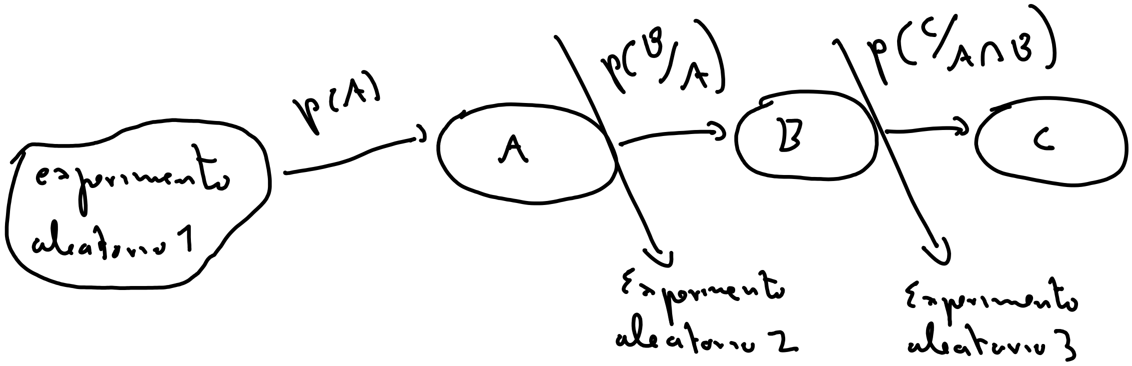


$$p\left(\text{persona enferma}/\text{test positivo}\right) = \frac{\frac{1}{10.000}\cdot\frac{99.99}{100}}{\frac{0.02}{100}\cdot\frac{9999}{10000} + \frac{99.99}{100}\cdot\frac{1}{10.000}} \approx \frac{1}{3}$$

* El concepto de probabilidad condicionada se puede extender al caso en el que tenemos la intersección de varios sucesos. P.ej.

$$p(A\cap B\cap C)=p(A\cap B)\cdot p(C/_{A\cap B})=p(A)\cdot p(B/_{A})\cdot p(C/_{A\cap B})$$

$$=p(B)\cdot p(A/_{B})\cdot p(C/_{A\cap B})=p(A\cap C)\cdot p(B/_{A\cap C})=p(A)\cdot p(C/_{A})\cdot p(B/_{A\cap C})$$



Para ir desde el origen hasta el C tenemos que ir multiplicando las probabilidades que nos encontramos por el camino

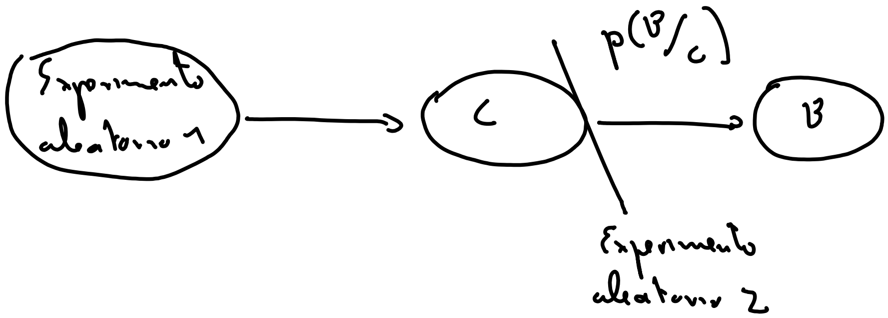
* Supongamos que B y C son dos sucesos y la unión de la familia $\{A_i\}$ de sucesos mutuamente excluyentes es todo el espacio muestral. Probar que

$$p(B/_{C})=\sum_i p(A_i/_{C})\,p(B/_{A_i\cap C})$$

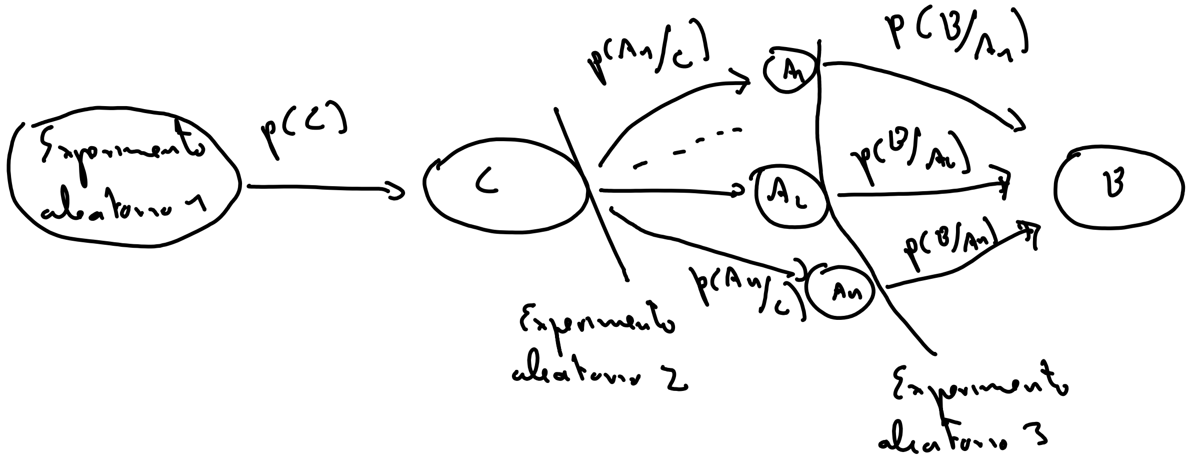
Prueba $$p(B\cap C)=p(B\cap C\cap E)=\sum_i p(A_i\cap B\cap C)=\sum_i p(B/_{A_i\cap C})\cdot p(A_i\cap C)=$$

$$=\sum_i p(B/_{A_i\cap C})\cdot p(A_i/_{C})\cdot p(C)$$

Pero $p(B\cap C)=p(C)\cdot p(B/_{C})$ Por tanto, combinando ambas fórmulas se llega al resultado buscado.



El camino que va de C a B se descompone en una suma de caminos que pasan por los sucesos A_i



COMBINATORIA

Recordemos la fórmula de Laplace encontrada para calcular la probabilidad de sucesos asociados a experimentos aleatorios con un espacio muestral finito:

$$p(A) = \frac{\text{casos favorables al suceso } A}{\text{casos posibles}} = \frac{\text{nº de sucesos elementales que constituyen } A}{\text{nº total de elementos del espacio muestral}}$$

Mirando esta fórmula vemos que necesitamos saber contar los elementos que se corresponden con cada situación. La combinatoria trata precisamente de eso, de la enumeración y el contaje de los subconjuntos de un conjunto finito de elementos.

Permutaciones

Imaginemos la siguiente situación: Tenemos un conjunto de n elementos diferentes y nos preguntamos de cuántas formas posibles se pueden ordenar:

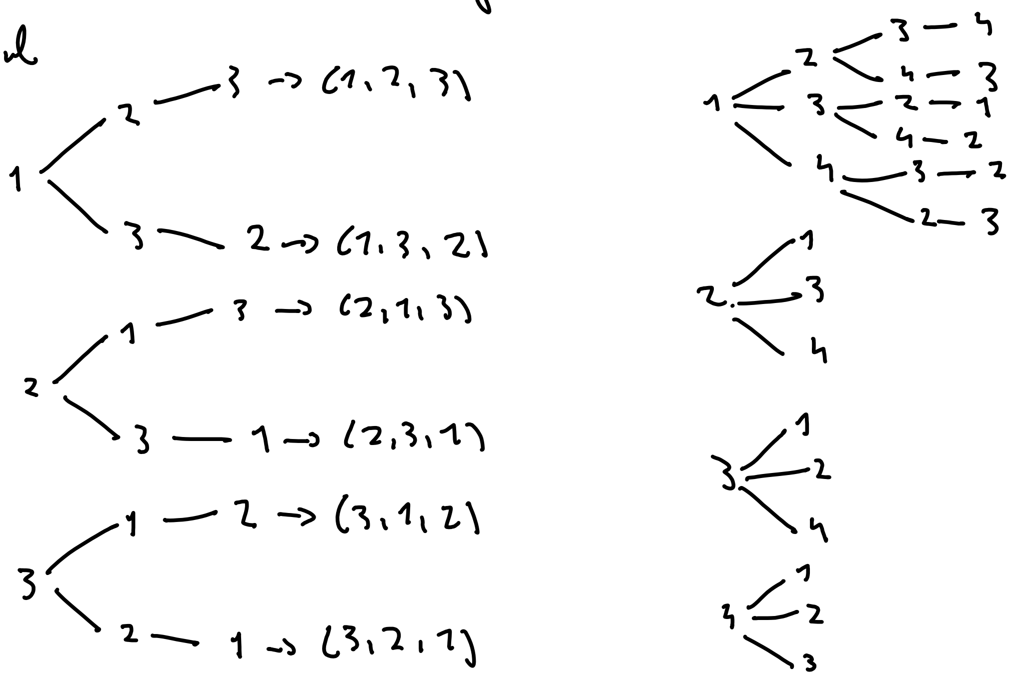
Ej. Tomamos el conjunto de los 3 primeros números naturales $(1,2,3)$. Se tienen las siguientes ordenaciones posibles

$(1,2,3)$ $(2,3,1)$

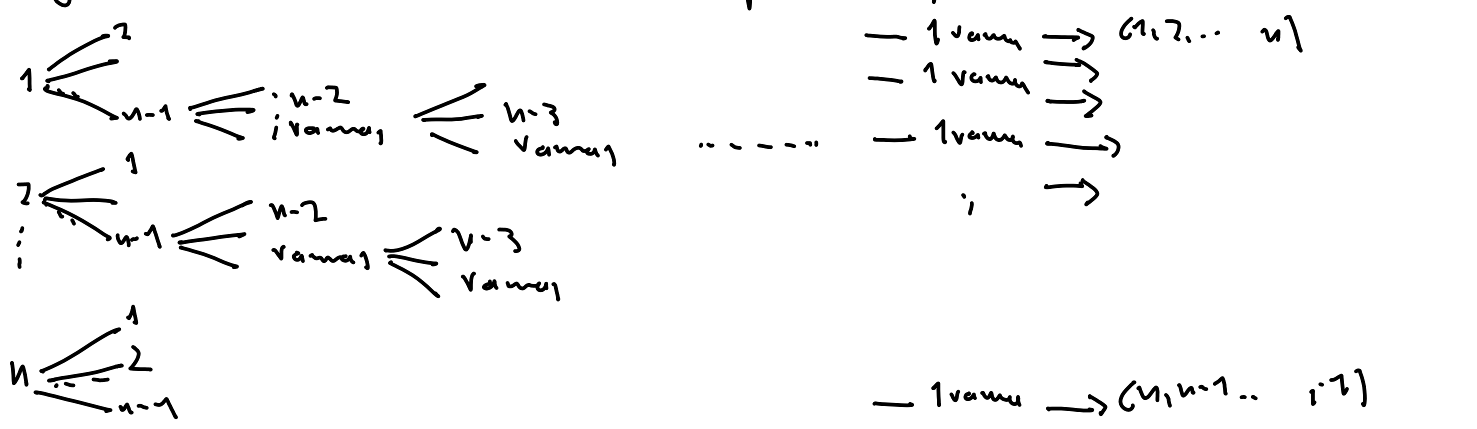
$(1,3,2)$ $(3,1,2)$

$(2,1,3)$ $(3,2,1)$

Cada una de las ordenaciones se denomina <u>permutación</u> del conjunto $(1,2,3)$. Podemos obtenerlas todas de forma sistemática a través del siguiente diagrama de árbol



El diagrama se generaliza sin dificultad al caso general: Tenemos n elementos distintos y queremos saber de cuántas maneras posibles se pueden ordenar. (nº de permutaciones)



Para saber el número total de permutaciones sólo tenemos que sumar el número total de ramas finales que viene dado por los productos del número de ramas que nos van saliendo en cada paso

$$n! = n\times(n-1)\times(n-2)\times(n-3)\times\cdots\times 1$$

Este es el número de permutaciones de un conjunto de n elementos denotado por P_n

Así que $P_n = n!$

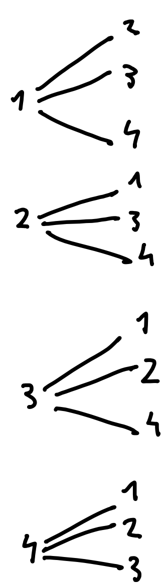
- A veces no nos interesa saber el número de permutaciones de TODOS los elementos de un conjunto dado, sólo necesitamos las permutaciones de un subconjunto de k elementos $k<n$

P^n_k = Número de permutaciones de n elementos tomados de k en k

Ejemplo P^4_2 = número de permutaciones de 4 elementos tomadas de 2 en 2

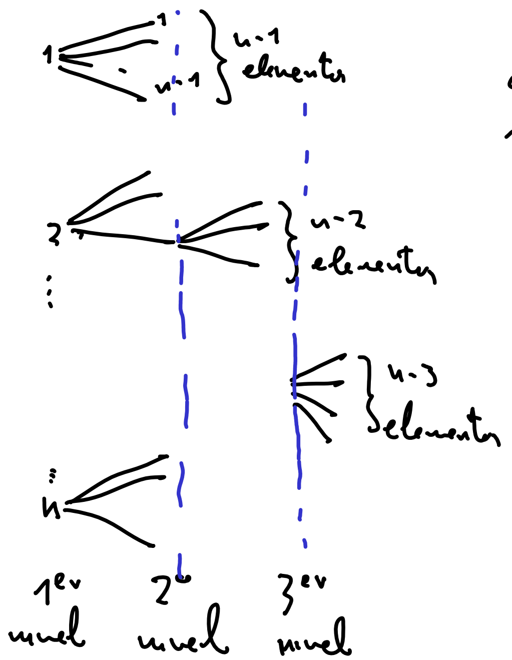
$$\left.\begin{matrix} (1\,2),\ (2\,1),\ (1\,3),\ (3\,1),\ (1\,4),\ (4\,1) \\ (2,3),\ (3,2),\ (2,4),\ (4,2) \\ (3,4),\ (4,3) \end{matrix}\right\}\ P^4_2 = 12$$

De nuevo podemos utilizar un diagrama de árbol para obtenerlas todas de forma sistemática.



$$P^4_2 = 4\times 3 = 12$$

Este ejemplo se genera sin dificultad para el caso general



El diagrama tiene tantos subniveles como elementos tienen los subconjuntos que estamos considerando

Por tanto

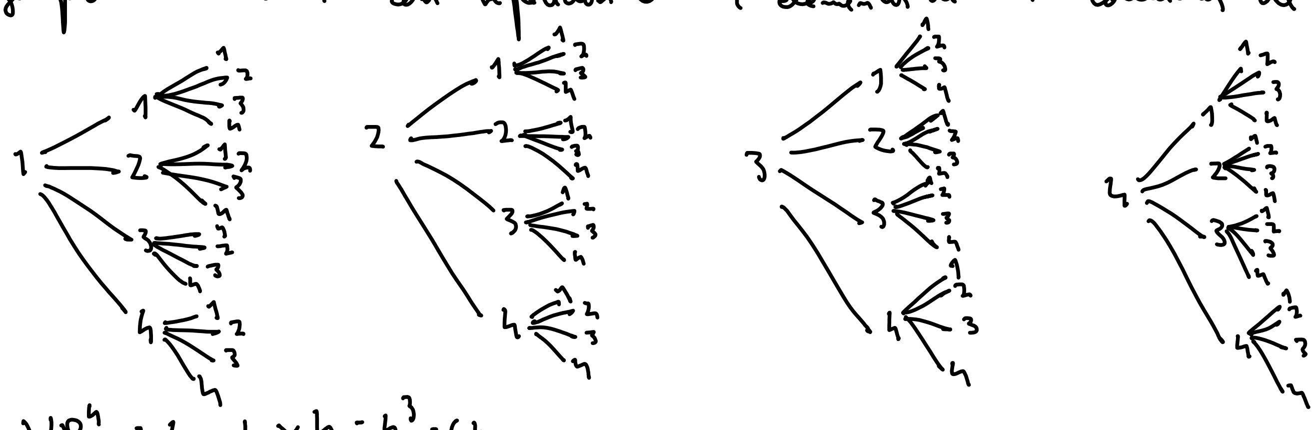
$$P^n_k = n\times(n-1)\times(n-2)\cdots\times(n-k+1) = \frac{n!}{(n-k)!}$$

* En algunos textos las permutaciones de n elementos tomados de k en k se conocen también como variaciones de n elementos tomados de k en k. Así que

$$P^n_k = V^n_k$$

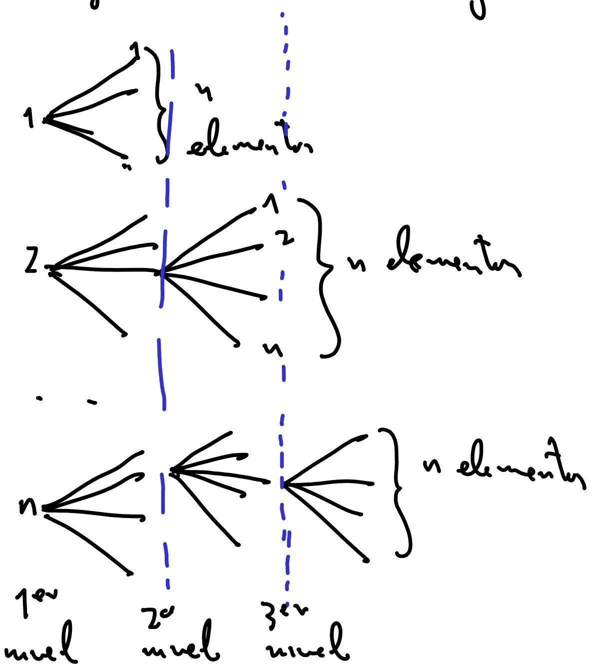
* Es posible generalizar lo anterior al caso en el que en las ordenaciones puedan aparecer elementos repetidos. Dos posibilidades: variaciones con repetición y permutaciones con repetición

* Ejemplo variaciones con repetición de 4 elementos distintos tomados de 3 en 3



$$VR^4_3 = 4\times 4\times 4 = 4^3 = 64$$

En general tendríamos el siguiente diagrama de árbol

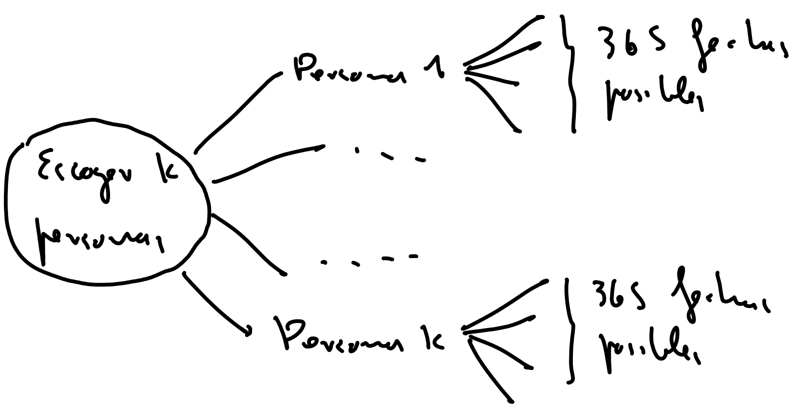


El diagrama tiene tantos subniveles como elementos tienen los subconjuntos que estamos considerando

$$VR^n_n = \overbrace{n \times n \times n \cdots \times n}^{n \text{ veces}} = n^n$$

→ Ejemplo: Calcular la probabilidad de que en un grupo de k personas escogidas al azar al menos 2 cumplan años el mismo día (excluído el 29 de Febrero)

Experimento aleatorio: escoger k personas al azar y comprobar su fecha de nacimiento



Es más fácil calcular la probabilidad del suceso complementario

$\bar{A}$: "Ninguna persona cumple años en la misma fecha" =

: "las fechas de los cumpleaños de las k personas son diferentes"

$$p(\bar{A}) = \frac{\text{nº de maneras de escoger k fechas diferentes (nos importa el orden porque las personas son diferentes)}}{\text{nº de maneras de escoger k fechas (nos importa el orden porque las personas son diferentes, y además las fechas pueden estar repetidas)}} = \frac{P^{365}_k}{VR^{365}_k} = \frac{365!}{(365-k)! \cdot 365^k}$$

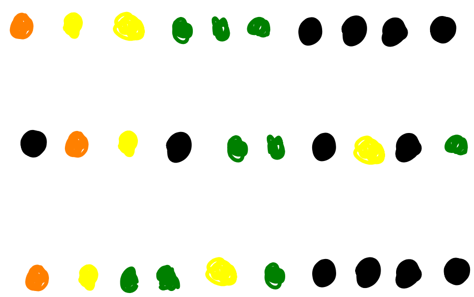
así que $p(A) = 1 - p(\bar{A}) =$

* Hasta ahora sólo hemos considerado ordenaciones en las que los elementos son distinguibles. Consideramos ahora ordenaciones con elementos que son indistinguibles de acuerdo con algún criterio.

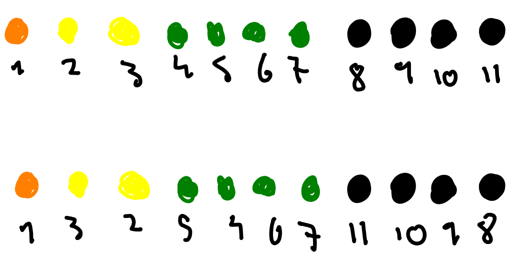
* Ejemplo: tenemos un conjunto de bolas de billar formado por una bola naranja, dos bolas amarillas, tres bolas verdes y cuatro bolas negras
¿De cuántas formas distinguibles se pueden ordenar las bolas?

Para responder al problema tenemos que tener en cuenta que las bolas del mismo color son indistinguibles:

Ejemplos de ordenaciones diferentes:



La forma de calcular el número total de ordenaciones es el siguiente: comenzamos suponiendo que las bolas son distintas, así que podemos numerar las bolas y después "quitamos" las ordenaciones que "sobran"



estas 2 ordenaciones serían diferentes si todas las bolas fuesen distinguibles.

tenemos que identificar las ordenaciones que son equivalentes debido a que las bolas son distinguibles,

Calculamos el número total de ordenaciones suponiendo que las bolas fuesen diferentes y lo vamos dividiendo por el número total de ordenaciones de las bolas de cada color para descontar los casos que tenemos de más

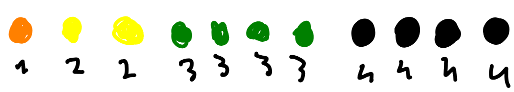
$$\frac{\text{Permutaciones de 15 elementos}}{\text{Permutaciones de 2 elementos (bolas amarillas)}} = \frac{15!}{2!}$$

$$\frac{\frac{15!}{2!}}{\substack{\text{Permutaciones}\\ \text{de 4 elementos}\\ \text{(bolas verdes)}}} = \frac{\frac{15!}{2!}}{4!} = \frac{15!}{2!\,4!} \qquad \frac{\frac{15!}{2!\,4!}}{\substack{\text{Permutaciones}\\ \text{de 4 elementos}\\ \text{(bolas negras)}}} = \frac{15!}{2!\,4!\,4!}$$

* En general: Tenemos n elementos de los cuales n_1 son indistinguibles entre sí, n_2 son indistinguibles entre sí, n_k son indistinguibles entre sí y $n_1 + \dots n_k = n$. Entonces el número de ordenaciones distintas de los n elementos viene dado por

$$PR^{n}_{n_1 n_2 \cdots n_k} = \frac{n!}{n_1!\, n_2! \cdots n_k!}$$

* Este número se conoce como las "permutaciones con repetición de n elementos con n_1 repeticiones, n_2 repeticiones ... , n_k repeticiones. Lo vemos en el ejemplo de las bolas



← La indistinguibilidad es equivalente a la repetición de elementos

* Hasta ahora hemos trabajado con ordenaciones en las que nos importa el orden de los elementos que estamos ordenando (permutaciones o variaciones). Ahora vamos a trabajar con un tipo de ordenaciones en las cuales el orden de los elementos es irrelevante. Estas ordenaciones se denominan combinaciones.

Gobsmackingly gorgeous

Ejemplo: Tenemos una pera, una manzana y una naranja y queremos preparar una bolsa con 2 frutas. ¿Cuántas bolsas diferentes se pueden preparar?

3 elementos distintos: {Pera, manzana, naranja}

Bolsas diferentes: {Pera, manzana}, {Pera, naranja}, {manzana, naranja}

No nos importa el orden {Pera, manzana} = {manzana, Pera}

En general tenemos n elementos distintos y queremos saber las ordenaciones de k elementos, sin importarnos el orden = combinaciones de n elementos tomados de k en k = C^n_k

Si nos importase el orden estaríamos hablando de las permutaciones de n elementos tomados de k en k

$$P^n_k = \frac{n!}{(n-k)!}$$ $\Rightarrow$ tenemos que "juntar" las ordenaciones que son equivalentes debido a que el orden es irrelevante. Como son ordenaciones de k elementos tenemos que dividir por k! (cada k! tenemos una única ordenación)

número combinatorio

$$C^n_k = \binom{n}{k} = \frac{P^n_k}{P_k} = \frac{n!}{k!\,(n-k)!} \quad 0 \leq k \leq n$$

Coeficiente binomial

* Como es bien sabido, el coeficiente binomial aparece en la fórmula del desarrollo del binomio

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$ $\Rightarrow$ por convenio $0! = 1$ así que

$$\binom{n}{0} = \frac{n!}{0!\,n!} = \frac{1}{0!} = 1$$

$$\binom{n}{n} = \frac{n!}{n!\,0!} = 1$$

* Ejemplo. Se baraja una baraja francesa y se reparten 13 cartas. ¿Cuál es la probabilidad de que haya 2 ases en las 13 cartas?

Experimento aleatorio $\Rightarrow$ se extraen 13 cartas de la baraja

$$p = \frac{\text{Nº de combinaciones de 13 cartas con 2 ases}}{\text{Nº total de combinaciones de 13 cartas}} = \frac{C^4_2 \times C^{48}_{11}}{C^{52}_{13}} = \frac{\frac{4!}{2!\,2!} \times \frac{48!}{11!\,37!}}{\frac{52!}{13!\,39!}} =$$

$$= \frac{48!\, 4!\, 13!\, 39!}{2!^2\, 11!\, 37!\, 52!} = \frac{39\cdot 38\cdot 13\cdot 12\cdot 4\cdot 3}{52\cdot 51\cdot 50\cdot 49\cdot 2} = \frac{4446}{20825} \approx 0.213493$$

* Ejemplo: Tenemos n objetos distinguibles y m cajas con capacidades $n_1, n_2 \dots n_m$, de tal forma que $n_1+n_2+\dots+n_m = n$. ¿De cuántas maneras posibles se pueden distribuir los objetos?

$$C^{n}_{n_1} \left\{ \begin{matrix} \cdot \to \left\{ C^{n-n_1}_{n_2} \to \left\{ C^{n-n_1-n_2}_{n_3} \quad \dots\dots \right. \right. \\ \vdots \\ \cdot \end{matrix} \right.$$

$C^{n}_{n_1}$: posibles maneras de meter los n objetos en la primera caja

$C^{n-n_1}_{n_2}$: posibles maneras de meter los objetos que quedan en la 2ª caja

$C^{n-n_1-n_2}_{n_3}$: posibles maneras de meter los objetos que quedan en la 3ª caja

Al final tenemos que:

$$C^{n}_{n_1} \times C^{n-n_1}_{n_2} \times C^{n-n_1-n_2}_{n_3} \times C^{n-n_1-n_2-n_3}_{n_4} \dots \times \underbrace{C^{\overbrace{n-n_1-n_2-\dots-n_{m-1}}^{n_m}}_{n_m}}_{C^{n_m}_{n_m}=1} =$$

$$= \frac{n!}{n_1!\,(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!\,(n-n_1-n_2)!} \cdot \frac{(n-n_1-n_2)!}{n_3!\,(n-n_1-n_2-n_3)!} \dots$$

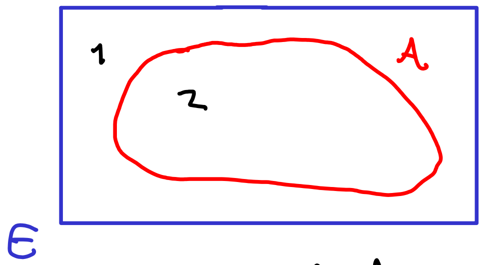
$$= \frac{n!}{n_1!\, n_2! \cdots n_m!} = PR^{n}_{n_1, n_2 \dots n_m} \Rightarrow \text{coeficientes multinomiales}$$

Recordamos que los coeficientes multinomiales aparecen en la expansión de $(x_1+\dots+x_m)^n$

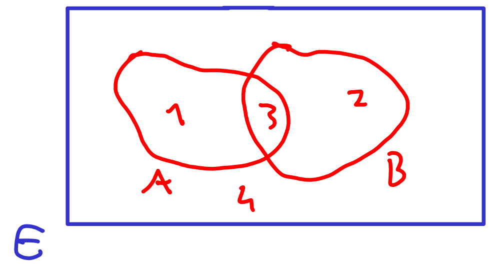
$$(x_1+\dots+x_m)^n = \sum_{\substack{n_1 \dots n_m \\ n_1+\dots+n_m=n}} \frac{n!}{n_1! \cdots n_m!}\, x^{n_1} \dots x^{n_m}$$

Ejemplo:

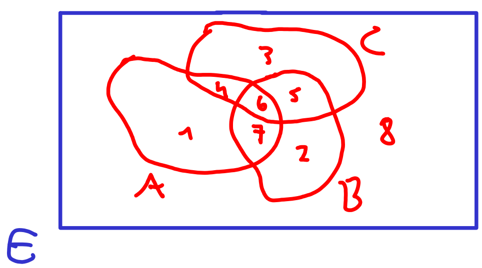
Si definimos sucesos, el espacio muestral E queda automáticamente dividido en un cierto número máximo de regiones. Lo vemos con los diagramas de Venn.



2 regiones (resultados que pertenecen al suceso A y resultados que no pertenecen al suceso A



4 regiones ya estudiadas



8 regiones

Cada suceso define 2 regiones $\rightarrow$ dentro / $\rightarrow$ fuera

Si tenemos n sucesos $\Rightarrow 2^n$ regiones

$\underbrace{\text{dentro dentro} \ldots \text{dentro}}_{n} \;\|\;\; \|\; \underbrace{\text{fuera fuera} \ldots \text{fuera}}_{n}$

Las 2^n regiones se pueden clasificar de la siguiente manera

- Región sin resultados 1.
- Regiones con resultados que sólo pertenecen a un suceso n
- Regiones con resultados que pertenecen a la intersección de 2 sucesos $\Rightarrow$ Combinaciones de n sucesos tomados de 2 en 2 $= C^n_2$
- Regiones con resultados que pertenecen a la intersección de 3 sucesos $\Rightarrow$ Combinaciones de 3 sucesos tomados de 3 en 3 $\Rightarrow C^n_3$

. . .

- Regiones con resultados que pertenecen a la intersección de n sucesos $\Rightarrow$ Combinaciones de n sucesos tomados de n en n $\Rightarrow C^n_n$

$$\sum_{k=0}^{n} C^n_k = 2^n$$

Ejemplo:

En el juego de "Bridge" se reparten las 52 cartas de una baraja francesa entre 4 jugadores. ¿Cuál es la probabilidad de que cada jugador reciba un as?

Cada persona recibe $\frac{52}{4} = 13$ cartas y como sólo hay 4 ases cada persona recibirá como máximo un as.

$$\begin{array}{r|l} 48 & 4 \\ \hline 08 & 12 \\ 0 & \end{array}$$

$$p = \frac{\text{nº de maneras diferentes que cada persona puede recibir un as}}{\text{nº total de formas de repartir 52 cartas distintas entre 4 personas}} = \frac{\text{nº de formas de repartir 52 cartas - 4 ases entre 4 personas} \times \text{Permutaciones de 4 ases} = 4!}{\frac{52!}{(13!)^4}} =$$

$$= \frac{\frac{48!}{(12!)^4}\,4!}{\frac{52!}{13!^4}} = \frac{4!\;48!\cdot 13!^4}{(12!)^4\cdot 52!} = \frac{24\;13^4}{52\cdot 51\cdot 50\cdot 49} \approx 0.105$$

* Combinaciones con repetición $\Rightarrow$

CR^n_k = combinaciones de n elementos distinguibles tomados de k en k de tal forma que es posible repetir elementos

Ejemplo: tenemos 4 elementos (1, 2, 3, 4) y queremos saber las combinaciones con repetición de estos elementos tomados de 2 en 2. Estas combinaciones vienen dadas por

(1,1), (1,2), (1,3), (1,4)

(2,2), (2,3), (2,4)

(3,3), (3,4)

(4,4)

Podemos utilizar la siguiente representación para estas combinaciones

(1, 1)	(1, 2)	(1, 3)	(1, 4)
xx|||	x|x||	x||x|	x|||x
(2, 2)	(2, 3)	(2, 4)	
|xx||	|x|x|	|x||x	
(3, 3)	(3, 4)		
||xx|	||x|x		
(4, 4)			
|||xx			

Podemos imaginar que las aspas son objetos indistinguibles y las líneas verticales denotan las paredes de cajas. Así que estas ordenaciones son equivalentes a las ordenaciones que resultan de meter objetos indistinguibles en cajas. Dado que las configuraciones

|xx|| y |||xx son distintas

entonces debemos considerar que las cajas son distinguibles

En resumidas cuentas:

Combinaciones con repetición de 4 elementos tomadas de 2 en dos = Número de maneras que se pueden distribuir 2 objetos indistinguibles en 4 cajas distinguibles

En general

Combinaciones con repetición de n elementos tomadas de k en k = Número de formas diferentes de distribuir k objetos indistinguibles en n cajas distinguibles (numeradas)

Podemos representar las n cajas con $n-1$ líneas verticales (separadores) y los h objetos indistinguibles con aspas. Así que el número de combinaciones con repetición resulta ser el número de permutaciones con repetición de $n-1+h$ objetos (aspas + separadores) de los que tenemos $n-1$ repetidos (los separadores) y h repetidos (las aspas)

$$\boxed{CR^n_h = PR^{n-1+h}_{n-1,h} = \frac{(n-1+h)!}{(n-1)!\,h!} = \binom{n-1+h}{h}}$$

Esta expresión aparece en el desarrollo en serie del binomio con exponente un entero negativo

$$(a+b)^{-n} = \sum_{h=0}^{\infty} \binom{n-1+h}{h} \left(\frac{-b}{a^n}\right)^h \qquad |a| > |b|$$

Ordenaciones de n elementos tomados de h en h

- Sin repetición o elementos indistinguibles
  - Nos importa el orden de los elementos
    - Variaciones (Permutaciones) $V^n_h = \frac{n!}{(n-h)!}$ $(n<h)$
    - Permutaciones $P_n = V^n_n = n!$
  - No nos importa el orden $\Rightarrow$ Combinaciones $C^n_h = \binom{n}{h} = \frac{n!}{h!\,(n-h)!}$
- Con repetición (elementos indistinguibles)
  - Nos importa el orden
    - Variaciones con repetición $VR^n_h = n^h$
    - Permutaciones con repetición $PR^n_{n_1 n_2 \dots n_k} = \frac{n!}{n_1! \cdots n_k!}$
  - No nos importa el orden $\Rightarrow$ Combinaciones con repetición $CR^n_h = \binom{n+h-1}{h}$

Ejemplo: el estándard IEEE 754 es un sistema utilizado por los ordenadores para representar números finitos. Un número x se representa de la siguiente manera

$$x = (-1)^s \times c \times b^q \quad \begin{cases} s = \text{signo} \\ c = \text{mantisa} \\ b = \text{base} \begin{cases} 2 \text{ (representación binaria)} \\ 10 \text{ (representación decimal)} \end{cases} \\ q = \text{exponente} \end{cases}$$

Ejemplo:

$x = -12.345 = (-1)^1 \times 12345 \times 10^{-3}$ // $x = 25 = (-1)^0 \times 0.25 \times 10^2$

Un número de "double precision" requiere 1 bit para el signo, 52 bits para la mantisa y 11 bits para el exponente (en total $1 + 52 + 11 = 64$ bits). ¿Cuántos números se pueden codificar con este formato? ¿Cuántos dígitos decimales tendrán los números así codificados?

Solución: con este formato la representación de un número en formato binario es:

$$x = (-1)^{\text{signo}} \times (1.b_0 b_1 b_2 \cdots b_{51})_2 \times 2^{e-1023}, \quad e = \{0, 1, \ldots, 2047\}, \quad b_i = \{0, 1\}$$

Así que tenemos $\underset{\text{mantisa}}{2^{52}} \times \underset{\text{signo}}{2} \times \underset{\text{exponente}}{2^{11}} = 2^{64}$ números

El número tiene 53 dígitos binarios $1, b_0, b_1, \ldots b_{51}$

$1 = 2^0$

$0.1 = \frac{1}{2}$

$0.101 = \frac{1}{2} + \frac{1}{2^3}$

$2^{-1023} = \frac{1}{2^{1023}} = \frac{5^{1023}}{10^{1023}}$

$0.01 = \frac{1}{2^2}$

$53 \log_{10} 2$

$0.10 = \frac{1}{2}$

$10^x = 2 \Rightarrow x = \log_{10} 2$

$0.11 = \frac{1}{2} + \frac{1}{2^2}$

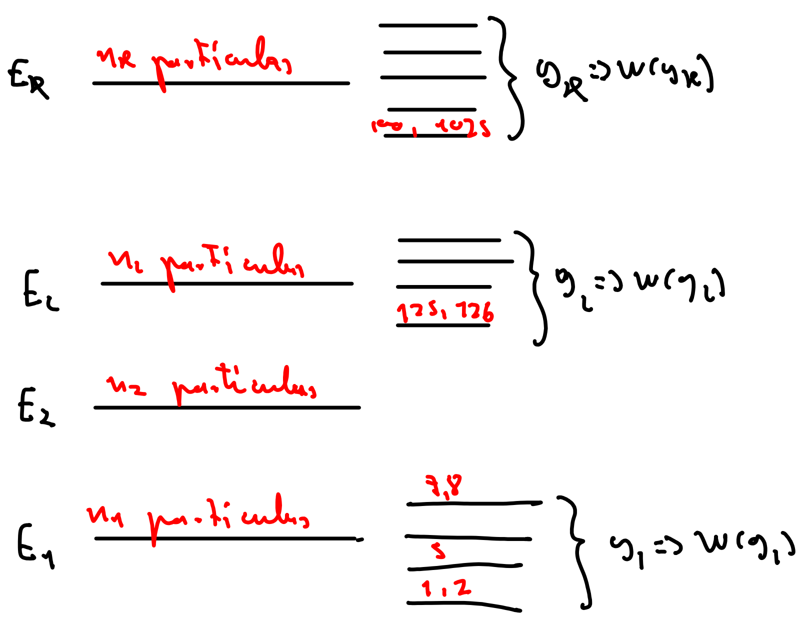
$10^{x53} = 2^{53} \Rightarrow x53 = 53 \log_{10} 2$

$0.111 = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3}$

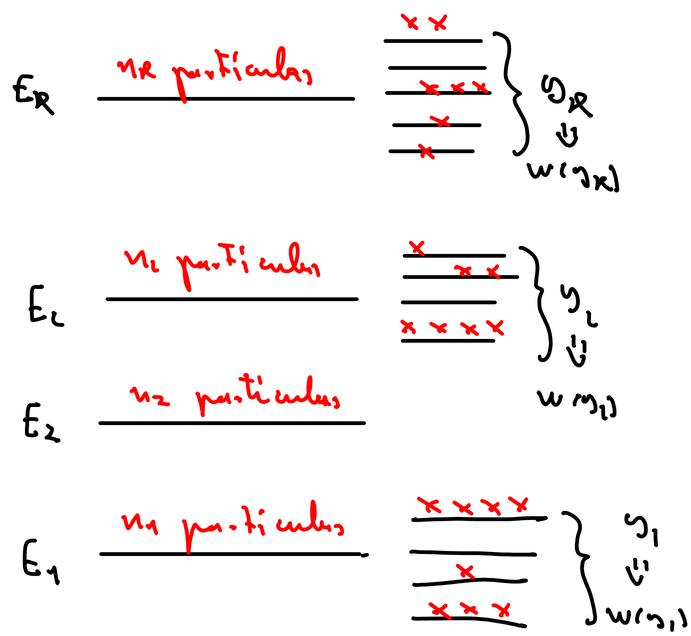
$10^A = 2^{53} \Rightarrow A = 53 \log_{10} 2$

• Ejemplo: Un sistema contiene N partículas que no interaccionan y cada una de ellas puede ocupar cualquiera de los estados cuánticos del sistema. Existen R niveles de energía $E_1, \ldots E_R$, y el nivel E_i posee una degeneración g_i (el nivel E_i tiene g_i estados cuánticos). Hallar el número de maneras distintas que las partículas pueden ocupar los estados cuánticos del sistema de tal forma que el nivel de energía i contenga n_i partículas, $i=1, \ldots R$ para cada uno de los siguientes casos:

i) Partículas distinguibles:



ii) Partículas indistinguibles



$w(g_i) \Rightarrow$ número de maneras diferentes de distribuir n_i partículas entre los subniveles g_i

$\Downarrow$

El problema es el mismo al de como distribuir N objetos diferentes entre R cajas con capacidades

$$n_1, n_2, \ldots n_R \Rightarrow \frac{N!}{n_1!\, n_2! \cdots n_R!}$$

Cada una de estas ordenaciones nos agrupa $w(g_1) \ldots w(g_R)$ ordenaciones diferentes en los subniveles así que

$$\Rightarrow \frac{N!}{n_1!\, n_2! \cdots n_R!}\, w(g_1) \cdots w(g_R)$$

$\Downarrow$

Aquí tenemos que ir multiplicando cada una de las diferentes maneras que tenemos de distribuir las partículas entre los subniveles

$$w(g_1) \cdots w(g_R)$$

Casos posibles :

i) Partículas distinguibles, sin restricción en el número de partículas que cada estado puede contener (Estadística de Maxwell-Boltzman)

niveles $g_i \to a, b, c, d$

partículas $\to$ 1, 2, 3 $\to$ asociar a cada partícula su nivel

1 2 3	1 2 3	1 2 3
a a a	a a b	d d d

1 2 3
b a a

$w(g_i) =$ Variaciones con repetición de g_i elementos tomados de n_i en n_i $: VR^{g_i}_{n_i} = g_i^{n_i}$

$$w(n_1 \dots n_k) = \frac{N!}{n_1!\, n_2! \cdots n_k!}\, g_1^{n_1} \cdots g_k^{n_k}$$

ii) Partículas indistinguibles, sin restricción en el número de partículas que cada estado puede contener (Estadística de Bose-Einstein)

niveles $g_i \to a, b, c, d$

partículas $\to$ x x x $\to$ asociar a cada partícula su nivel.

$\to$ Es lo mismo que antes pero ahora no nos importa el orden porque las partículas son distinguibles.

x x x	x x x	=	x x x
a a a	a a d	=	d a a

$$w(g_i) = CR^{g_i}_{n_i} = \binom{n_i + g_i - 1}{n_i}$$

iii) Partículas indistinguibles con un máximo de una partícula por estado (Estadística de Fermi-Dirac)

niveles $g_i \to a, b, c, d \to$ asociar a cada partícula su nivel

Partículas x x x

$\to$ Es lo mismo que el caso anterior pero ahora no se pueden repetir elementos

x x x	x x x	:	x x x
a b c	c d b	:	b d c

$$w(g_i) = C^{g_i}_{n_i} = \binom{g_i}{n_i} = \frac{g_i!}{n_i!\,(g_i - n_i)!}$$

iv) Partículas distinguibles, con un máximo de una partícula por estado

niveles $g_i \to a, b, c, d$

partículas $\to$ 1, 2, 3 $\to$ asociar a cada partícula su nivel

Es lo mismo que el caso anterior pero ahora nos importa el orden:

1 2 3	1 2 3	1 2 3
a b c	c a b	b d a

$$w(g_i) = V^{g_i}_{n_i} = \frac{g_i!}{(g_i - n_i)!}$$

Variables aleatorias y distribuciones de probabilidad

Tenemos un experimento aleatorio, y su espacio muestral asociado E. Podemos asociar un número a cada elemento del espacio muestral. Si el espacio muestral es finito o numerable entonces el número que asociamos a cada elemento suele ser un entero y si es infinito y no numerable, le asociaremos un número real o varios.

Experimentos aleatorios $\Rightarrow$

(1) Láncese un dado de seis caras $\Rightarrow E = \{1, 2, 3, 4, 5, 6\}$

(2) Láncese una moneda $\Rightarrow E = \{+, -\}$

(3) Colóquese un radar en una carretera y mídase la velocidad de los coches $\Rightarrow E = (0 \, km/h, 350 \, km/h)$

Vemos que en el primer y en el tercer experimento aleatorio ya hay una forma natural de asignar un número a cada elemento del espacio muestral. En el segundo experimento lo tenemos que hacer nosotros.

Sale $+ \rightarrow$ asignamos el 0. Sale $- \longrightarrow$ asignamos el 1.

Definición: una variable aleatoria es una aplicación entre el espacio muestral de un experimento aleatorio y un (sub)-conjunto I de los números reales. El dominio de la aplicación es todo el espacio muestral. Si el conjunto I es finito o es numerable entonces se dice que la variable aleatoria es discreta. En caso contrario se habla de una variable aleatoria continua. I es el recorrido de la variable aleatoria.

Por tanto podemos utilizar los números reales que hemos usado para definir la variable aleatoria para designar a cada uno de los elementos del espacio muestral.

En los ejemplos (1) y (2) definen variables aleatorias discretas y (3) define una variable aleatoria continua.

* Notación para las variables aleatorias: usamos una letra + dos puntos + el recorrido de la variable (conjunto discreto/finito o conjunto infinito)

Ejemplo (1) $\Rightarrow X : \{1, 2, 3, 4, 5, 6\}$

Ejemplo (2) $\Rightarrow X : \{0, 1\}$

Ejemplo (3) $\Rightarrow X : (0, 350)$

$\Rightarrow$ A veces utilizaremos sólo la letra si el recorrido de la variable aleatoria está claro por el contexto o no hace falta especificarlo.

* Variables aleatorias discretas: Sean $x_1, \dots, x_n$ los números reales utilizados en la definición de una variable aleatoria discreta (supondremos que $\{x_1, \dots x_n\}$ es un conjunto finito, pero la generalización para el caso infinito es fácil). Entonces la probabilidad de cada uno de los resultados que definen el espacio muestral se puede representar a través de una función f con dominio $\{x_1, \dots x_n\}$ y recorrido $[0, 1]$

$$f : \{x_1, \dots x_n\} \longrightarrow [0, 1]$$

$$f(x_i) \longmapsto p(x_i) = \text{probabilidad del suceso que se corresponde con el número } x_i$$

La función f se denomina función de probabilidad. Claramente

$$\sum_{i=1}^{n} f(x_i) = 1$$

A partir de la función de probabilidad f podemos definir la función de probabilidad acumulada F del siguiente modo:

$$F : \{x_1, \dots, x_n\} \longrightarrow [0, 1]$$

$$F(x_i) \longmapsto \sum_{k=1}^{i} f(x_k)$$

*Ejemplo: una bolsa contiene siete bolas rojas y tres bolas blancas. Se extraen tres bolas sin devolución. Hallar la función de probabilidad de las bolas rojas.

Experimento aleatorio $\Rightarrow$ extracción de tres bolas de una bolsa

Espacio muestral $\Rightarrow \{\bullet\bullet\bullet\}, \{\bullet\bullet\circ\}, \{\bullet\circ\circ\}, \{\circ\circ\circ\}$

$$\begin{array}{cccc} \{\bullet\bullet\bullet\} & \{\bullet\bullet\circ\} & \{\bullet\circ\circ\} & \{\circ\circ\circ\} \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 2 & 1 & 0 \end{array}$$

Variable aleatoria $\Rightarrow X$ (nº de bolas rojas) $\Leftarrow$ valores de la variable aleatoria.

función de probabilidad $\Rightarrow$
$$p(3) = \frac{\text{nº de ordenaciones con tres bolas rojas}}{\text{nº de ordenaciones con 3 bolas}} = \frac{C^7_3}{C^{10}_3} = \frac{\frac{7!}{3!\,4!}}{\frac{10!}{3!\,7!}} = \frac{3!\,7!\cdot 7!}{3!\,4!\cdot 10!} = \frac{7\cdot 6\cdot 5}{10\cdot 9\cdot 8} = \frac{7}{24}$$

$$p(2) = \frac{\text{nº de ordenaciones con 2 bolas rojas}}{\text{nº de ordenaciones con 3 bolas}} = \frac{C^7_2 \times 3}{C^{10}_3} = \frac{\frac{7!}{2!\,5!}\times 3}{\frac{10!}{3!\,7!}} = \frac{7!\;3\;3!\;7!}{2!\;5!\;10!} = \frac{7\cdot 6\cdot 3\cdot 3}{10\cdot 9\cdot 8} = \frac{21}{40}$$

$$p(1) = \frac{\text{nº de ordenaciones con 1 bola roja}}{\text{nº de ordenaciones con 3 bolas}} = \frac{7\times C^3_2}{C^{10}_3} = \frac{7\times\frac{3!}{2!}}{\frac{10!}{3!\,7!}} = 7\,\frac{3!\;3!\;7!}{2!\;10!} = 7\,\frac{3\cdot 6}{10\cdot 9\cdot 8} = \frac{7}{40}$$

$$p(0) = \frac{\text{nº de ordenaciones con 0 bolas rojas}}{\text{nº de ordenaciones con 3 bolas}} = \frac{1}{C^{10}_3} = \frac{3!\,7!}{10!} = \frac{3\cdot 2}{10\cdot 9\cdot 8} = \frac{1}{120}$$

Nótese que $p(0)+p(1)+p(2)+p(3) = \frac{1}{120}+\frac{7}{40}+\frac{21}{40}+\frac{7}{24} = 1$ $\qquad \begin{array}{r|l} 64 & 16 \\ 00 & 4 \end{array}$

* Variables aleatorias continuas

Si la imagen de la biyección que utilizamos para definir la variable aleatoria es el conjunto de los números reales $\mathbb{R}$ (o un subconjunto de la recta real que se pueda poner en biyección con $\mathbb{R}$) entonces diremos que la variable aleatoria es continua:

Ejemplos:

Experimentos aleatorios	Variables aleatorias continuas
Colóquese un radar en una carretera y mídase la velocidad de los coches	$X: (0\, km/h, 400\, km/h)$
Escójase una persona al azar y mídase su altura	$X: (0\, kg, 200\, kg)$

Para fijar ideas consideraremos una variable aleatoria $X: (p, q)$ donde $I = (p, q)$ es un intervalo de la recta real. La función densidad de probabilidad asociada a la variable aleatoria es una función $f: I \to \mathbb{R}^+$ tal que la probabilidad del suceso $A \subset I$ viene dada por

$$p(A) = \int_A f(x)\, dx \quad .$$

p. ej. $A = [a, b] \subset (p, q)$ entonces $p(A) = \int_a^b f(x)\, dx$

Si tenemos el suceso $E = [p,q]$ entonces la condición $p(E)=1$ implica que $\int_p^q f(x)\,dx = 1$.

Muchas veces se toma $E=(-\infty,\infty)$ por lo que en ese caso

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

* La función de probabilidad acumulada $F(x)$ asociada a la función densidad de probabilidad $f(x)$ se define así

$$F(x) = \int_p^x f(u)\,du$$

* Así que los Teoremas básicos del cálculo nos dicen que

$$p([a,b]) = \int_a^b f(x)\,dx = F(b) - F(a) \qquad \frac{dF(x)}{dx} = f(x)$$

* Ejemplo: una variable aleatoria tiene la siguiente función densidad de probabilidad

$$f(x) = \begin{cases} Ae^{-x} & 0 < x < \infty \\ 0 & \infty < x < 0 \end{cases}$$

Encuéntrese el valor de la constante A y hállese la probabilidad del suceso $(1,2]$

La constante A se halla utilizando la condición de normalización

$$1 = \int_0^{\infty} Ae^{-x}\,dx = \left[-Ae^{-x}\right]_0^{\infty} = A$$

Para calcular la probabilidad podemos usar la definición de función densidad de probabilidad.

$$p(1,2] = \int_1^2 dx\, e^{-x} = \left[-e^{-x}\right]_1^2 = -e^{-2} + e^{-1}$$

* Utilizaremos la terminología "Distribución de probabilidad" para referirnos a una función de probabilidad (variable aleatoria discreta) o una función densidad de probabilidad (variable aleatoria continua). => comprobar terminología.

* La función densidad de probabilidad no aporta la suficiente información para calcular la probabilidad de sucesos individuales => ¿Verdadero o falso?

* Si tenemos una función de probabilidad p definida sobre una variable aleatoria discreta $x: \{x_1, .., x_n\}$ entonces podemos definir a partir de ella una función densidad de probabilidad f definida sobre una variable aleatoria continua $x: (-\infty, \infty)$ utilizando la "delta de Dirac".

$$f(x) = \sum_{i=1}^{n} p(x_i)\, \delta(x - x_i) \Rightarrow \text{si } I = (a,b) \text{ y } a \leq x_1 < x_2 \dots < x_n \leq b \text{ entonces:}$$

$$= p((a,b)) = \int_a^b f(x)\,dx = \int_a^b \sum_{i=1}^{n} p(x_i)\,\delta(x - x_i)\,dx = \sum_{i=1}^{n} p(x_i) \int_a^b \delta(x - x_i)\,dx = \sum_{i=1}^{n} p(x_i)$$

Si el intervalo (a,b) no contiene todos los valores $x_1, \dots x_n$ entonces la suma final sólo se extenderá a los puntos contenidos en (a,b)

* Conjunto de variables aleatorias => variables aleatorias multidimensionales. Podemos definir variables aleatorias multidimensionales sin más que efectuar el producto cartesiano de variables aleatorias unidimensionales. En general este tipo de variables aleatorias están relacionadas con conjuntos de más de un experimento aleatorio.

Ejemplo: Escogemos al azar una persona y medimos su altura y su peso

Variable aleatoria para la altura $\Rightarrow x: (0\,m, 3\,m)$

Variable aleatoria para el peso $\Rightarrow y: (0\,kg, 200\,kg)$

Producto cartesiano de variables aleatorias continuas $(x,y): (0\,m, 3\,m) \times (0\,kg, 200\,kg)$

Ejemplo: Tiramos dos dados y anotamos el valor de cada uno de los números que salen

Variable aleatoria para los resultados del 1er dado. $\Rightarrow X : \{1, 2, 3, 4, 5, 6\}$

Variable aleatoria para los resultados del 2º dado. $\Rightarrow y : \{1, 2, 3, 4, 5, 6\}$

Variable aleatoria para el experimento de los 2 dados: $(x, y) : \begin{Bmatrix} (1,1), (1,2), (1,3), \ldots & (1,6) \\ (2,1), (2,2), (2,3) \ldots & (2,6) \\ \vdots & \\ (6,1), (6,2), (6,3), \ldots & (6,6) \end{Bmatrix}$

* Los ejemplos se generalizan sin dificultad para casos de variables aleatorias con más "dimensiones" o argumentos

* Es posible definir funciones de probabilidad y funciones densidad de probabilidad para variables aleatorias multidimensionales. Las funciones así resultantes se denominan funciones de probabilidad o de densidad de probabilidad conjuntas.

Variable aleatoria discreta multidimensional $(x, y) : \{(x_i, y_j)\} \Rightarrow$ función de probabilidad $f(x_i, y_j) = p[(x_i, y_j)]$

Condición de normalización $1 = \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i, y_j)$

Variable aleatoria continua multidimensional $(x, y) : (-\infty, \infty) \times (-\infty, \infty) \Rightarrow$ función densidad de probabilidad $f(x, y)$

Condición de normalización $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$

* Dos variables aleatorias X, Y son independientes si los experimentos aleatorios asociados a las mismas son independientes. En este caso es posible definir una función (densidad) de probabilidad conjunta que es el producto de las funciones (densidad) de probabilidad de las variables X e Y

$$x:(-\infty,\infty)\ ,\quad y:(-\infty,\infty) \Rightarrow (x,y):(-\infty,\infty)\times(-\infty,\infty)$$

$$\Downarrow \qquad\qquad \Downarrow \qquad\qquad \Downarrow$$

$$f(x) \qquad\qquad g(y) \qquad\qquad \psi(x,y)=f(x)\,g(y)$$

*

Ejemplo: Tiramos dos dados y anotamos el valor de cada uno de los números que salen

Variable aleatoria para los resultados del 2º dado. $\Rightarrow y:\{1,2,3,4,5,6\}$

Variable aleatoria para los resultados del 1er dado. $\Rightarrow X:\{\underset{x_1}{1},\underset{x_2}{2},3,4,5,6\}$

Función de probabilidad $f(y_i)=\dfrac{1}{6}\quad \forall i=1,2,\dots,6$

Función de probabilidad $g(x_i)=\dfrac{1}{6}\quad \forall i=1,2..6$

Los experimentos aleatorios asociados a cada variable aleatoria son independientes. Así que si definimos la variable aleatoria (X,Y) entonces podemos definir una función de probabilidad conjunta así:

$$\psi(x,y)=f(x)\,g(y)=\frac{1}{36}\quad \left\{\begin{array}{l} \psi(1,1)=\frac{1}{36},\ \psi(1,2)=\frac{1}{36}\ \dots\ ,\ \psi(1,6)=\frac{1}{36} \\ \vdots \\ \psi(6,1)=\frac{1}{36},\ \psi(6,2)=\frac{1}{36}\ \dots\ ,\ \psi(6,6)=\frac{1}{36} \end{array}\right.$$

Esto es reminiscente de la definición de probabilidad de sucesos independientes (asociados a uno o varios experimentos aleatorios)

$$p(A\cap B)=p(A)\cdot p(B)$$

* Ejemplo: las variables aleatorias independientes X, Y tienen las respectivas funciones de densidad de probabilidad $g(x) = e^{-x}$ y $h(y) = 2e^{-2y}$. Calcúlese la probabilidad del suceso $1 < x \leq 2$ y $0 < y \leq 1$

$$\int_0^1 dy \int_1^2 2e^{-x} e^{-2y} dx = \int_0^1 dy\, e^{-2y} \int_1^2 2 dx\, e^{-x} = \frac{e^{-2y}}{-2}\Big|_0^1 \frac{2e^{-x}}{-1}\Big|_1^2 =$$

$$= (1 - e^{-2})(-e^{-2} + e^{-1}) = \frac{e^2 - 1}{e^2} \frac{e - 1}{e^2} = \frac{(e-1)^2(e+1)}{e^4} \simeq 0.201$$

Propiedades de las distribuciones de probabilidad

* Dada una variable aleatoria X, una distribución de probabilidad nos da normalmente toda la información que necesitamos. Sin embargo podemos definir una serie de parámetros que hasta cierto punto permiten caracterizar la distribución de probabilidad

Sea X una variable aleatoria y $g: X \to \mathbb{R}$ una función. Definimos el valor esperado de g $E(g(x))$ respecto a una distribución de probabilidad $f(x)$ asociada a X así

$$E(g(x)) = \begin{cases} \sum_i g(x_i) f(x_i) \Rightarrow \text{variable aleatoria discreta} \Rightarrow \text{serie absolutamente convergente} \\ \int f(x) g(x) dx \Rightarrow \text{variable aleatoria continua} \Rightarrow \text{La integral existe.} \end{cases}$$

* Una serie $\sum_{k=1}^{n} a_k$ es absolutamente convergente si $S_n = \sum_{k=1}^{n} |a_k|$ converge

* Propiedades elementales del valor esperado:

Si a es una constante entonces $E(a) = a$

Si a es una constante entonces $E(a g(x)) = a E(g(x))$

Si definimos $g(x) = s(x) + t(x)$ entonces $E(g(x)) = E(s(x)) + E(t(x))$

Así que el valor esperado E permite definir un operador lineal en el conjunto de funciones de variable aleatoria

* Media de una distribución de probabilidad ⇒ valor esperado de la propia variable aleatoria:

$$\langle x\rangle=\mu=E(x)=\begin{cases}\text{variable aleatoria discreta } \sum_i x_i f(x_i)\\ \text{variable aleatoria continua } \int x f(x)dx\end{cases}$$

► *The probability of finding a* 1s *electron in a hydrogen atom in a given infinitesimal volume* dV *is* $\psi^*\psi\, dV$, *where the quantum mechanical wavefunction* ψ *is given by*

$$\psi = Ae^{-r/a_0}.$$

Find the value of the real constant A *and thereby deduce the mean distance of the electron from the origin.*

Solución ⇒ función densidad de probabilidad $\psi\psi^* = A^2 e^{-\frac{2r}{a_0}}$

Elemento de volumen en coordenadas esféricas $dV = r^2 \sin\theta\, dr\, d\theta\, d\varphi$

La probabilidad de hallar al electrón en cualquier lugar del espacio debe ser igual a la unidad

$$1=\int_{\mathbb{R}^3} A^2 e^{-\frac{2r}{a_0}}dV=\int_{\mathbb{R}^3} A^2 e^{-\frac{2r}{a_0}} r^2\sin\theta\, dr\, d\theta\, d\varphi=\int_0^{2\pi}d\varphi\int_0^{\pi}d\theta \sin\theta\int_0^{\infty}A^2 r^2 e^{-\frac{2r}{a_0}}=$$

$$=2\pi A^2\left[-\cos\theta\right]_0^{\pi}\int_0^{\infty} r^2 e^{-\frac{2r}{a_0}}dr=2\pi A^2\, 2\,\frac{a_0^3}{4}=\pi A^2 a_0^3 \Rightarrow A=\frac{1}{a_0^{3/2}\sqrt{\pi}}=\frac{1}{(a_0^3\pi)^{\frac{1}{2}}}$$

$$\int r^2 e^{-\frac{2r}{a_0}}=\left[r^2\frac{e^{-\frac{2r}{a_0}}}{-\frac{2}{a_0}}+\int\frac{a_0}{2}2r e^{-\frac{2r}{a_0}}dr\right]=-\frac{a_0 r^2 e^{-\frac{2r}{a_0}}}{2}+a_0 r\frac{e^{-\frac{2r}{a_0}}}{-\frac{2}{a_0}}+\int\frac{a_0^2}{2}e^{-\frac{2r}{a_0}}=$$

$$=-\frac{a_0^2}{2}r e^{-\frac{2r}{a_0}}+\frac{a_0^3}{2}\frac{e^{-\frac{2r}{a_0}}}{-\frac{2}{a_0}}=-\frac{a_0^2 r}{2}e^{-\frac{2r}{a_0}}-\frac{a_0^3}{4}e^{-\frac{2r}{a_0}}=\int r^2 e^{-\frac{2r}{a_0}}$$

Distancia media del electrón al origen $\langle r\rangle=\int_{\mathbb{R}^3} r A^2 e^{-\frac{2r}{a_0}}dV=A^2\int_{\mathbb{R}^3} r e^{-\frac{2r}{a_0}} r^2\sin\theta\, dr\, d\theta\, d\varphi=$

$$=A^2\int_0^{2\pi}d\varphi\int_0^{\pi}\sin\theta\, d\theta\int_0^{\infty}r^3 e^{-\frac{2r}{a_0}}dr=2\pi A^2\left[-\cos\theta\right]_0^{\pi}3\frac{a_0^4}{8}=3\,\frac{4\pi A^2 a_0^4}{8}=\frac{3\pi\, a_0^4}{2\pi\, a_0^3}=\frac{3a_0}{2}$$

$$\int r^3 e^{-\frac{2r}{a_0}}dr = \left[\frac{r^3 e^{-\frac{2r}{a_0}}}{-\frac{2}{a_0}} + \frac{a_0}{2}\int 3r^2 e^{-\frac{2r}{a_0}}dr \right] = -\frac{a_0}{2}r^3 e^{-\frac{2r}{a_0}} - \frac{3a_0}{2}\left(\frac{a_0^2 r}{2} e^{-\frac{2r}{a_0}} + \frac{a_0^3}{4} e^{-\frac{2r}{a_0}} \right)$$

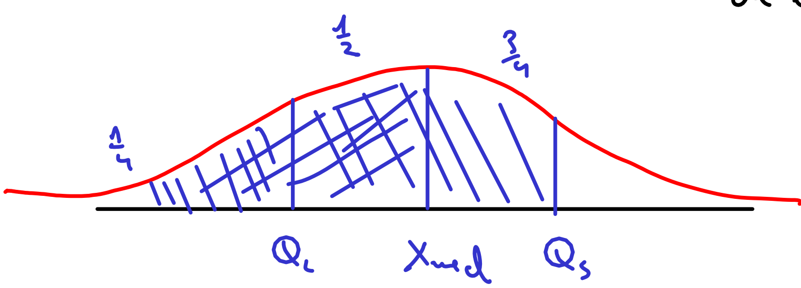
$$\Rightarrow \int_0^\infty r^3 e^{-\frac{2r}{a_0}}dr = \frac{3a_0^4}{8}$$

* Moda de una distribución de probabilidad $f(x) \Rightarrow$ es el valor x_{mod} de la variable aleatoria en el cual $f(x)$ alcanza el valor máximo. El valor x_{mod} no tiene por qué ser único en cuyo caso cualquiera de los valores posibles se denominará moda de la distribución de probabilidad

* Mediana de una distribución de probabilidad $f(x) \Rightarrow$ es el valor X_{med} de la variable aleatoria en el cual la función de probabilidad acumulada $F(x)$ alcanza el valor $\frac{1}{2}$, es decir, $F(X_{Med}) = \frac{1}{2}$.

$$\left.\begin{array}{l} \text{* Cuartil inferior } Q_i \Leftrightarrow F(Q_i) = \frac{1}{4} \\ \text{* Cuartil superior } Q_s \Leftrightarrow F(Q_s) = \frac{3}{4} \end{array}\right\}$$

La mediana, el cuartil inferior y el cuartil superior dividen la variable aleatoria en 4 regiones, conteniendo cada región $\frac{1}{4}$ de la probabilidad



* Percentil n-ésimo $P_n \Leftrightarrow F(P_n) = \frac{n}{100}$

* Varianza de una distribución de probabilidad $f(x)$ Si μ es el valor medio entonces

$$V(x) = \sigma^2 = E((x-E(x))^2) = E((x-\mu)^2) = \begin{cases} \sum_i (x_i-\mu)^2 f(x_i) \Rightarrow \text{variable aleatoria discreta} \\ \int (x-\mu)^2 f(x)dx \Rightarrow \text{variable aleatoria continua} \end{cases}$$

$V(x) \geq 0$

Si alguno de los valores anteriores no existe entonces la distribución no tiene varianza

* Desviación típica de una distribución de probabilidad $\sigma = (V(x))^{\frac{1}{2}}$.
  La desviación típica nos indica lo concentrada que está en torno a la media los valores que toma la distribución de probabilidad
* Propiedades

1 $V(a) = 0 \quad V(a) = E((a - E(a))^2) = E((a-a)^2) = 0$

2 $V(ax) = E((ax - E(ax))^2) = E((ax - aE(x))^2) = E(a^2(x - E(x))) = a^2 E(x - E(x)) = a^2 V(x)$

3 $V(ax+b) = E\left[(ax+b - E(ax+b))^2\right] = E\left[(ax+b - aE(x) - b)^2\right] = E\left[a^2(x - E(x))^2\right] =$
$= a^2 E\left[(x - E(x))^2\right] = a^2 V(x)$

►*Find the standard deviation of the PDF for the distance from the origin of the electron whose wavefunction was discussed in the previous two examples.*

Función densidad de probabilidad $\quad \psi\psi^* = A^2 e^{-\frac{2r}{a_0}} \qquad A = \frac{1}{a_0^{3/2}\sqrt{\pi}} = \frac{1}{(a_0^3\pi)^{\frac{1}{2}}} \Rightarrow$

Función densidad de probabilidad radial $\quad \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta \, r^2 \psi\psi^* = 4\pi r^2 \psi\psi^* = \frac{4\pi r^2}{a_0^3 \pi} e^{-\frac{2r}{a_0}} = \frac{4r^2 e^{-\frac{2r}{a_0}}}{a_0^3}$

$\psi\psi^* = \frac{1}{a_0^3\pi} e^{-\frac{2r}{a_0}} \Rightarrow$

Distancia media $= E(r) = \frac{3}{2} a_0$

Varianza: $E\left[(r - \frac{3}{2}a_0)^2\right] = \int_0^{\infty} dr \, (r - \frac{3}{2}a_0)^2 \frac{4r^2 e^{-\frac{2r}{a_0}}}{a_0^3} = \int_0^{\infty} dr \left(r^2 + \frac{9}{4}a_0^2 - 3r a_0\right) \frac{4r^2 e^{-\frac{2r}{a_0}}}{a_0^3} =$

$= \int_0^{\infty} dr \frac{4r^4}{a_0^3} e^{-\frac{2r}{a_0}} + \frac{9}{a_0} \int_0^{\infty} r^2 e^{-\frac{2r}{a_0}} dr - \frac{12}{a_0^2} \int_0^{\infty} r^3 e^{-\frac{2r}{a_0}} dr = \frac{3}{4} a_0^2$

Desviación típica $\Rightarrow \sigma = \left(\frac{3}{4} a_0^2\right)^{\frac{1}{2}} = \frac{a_0}{2}\sqrt{3}$.

* Desigualdad de Bienaymé-Chebyshev Bienaymé–Chebyshev

Para una distribución de probabilidad y un número $c>0$ se tiene que

$$p(|x-\mu|\geq c)\leq \frac{\sigma^2}{c^2}$$

$\mu = E(x) \Rightarrow$ media de la distribución de probabilidad

$\sigma = V(x) \Rightarrow$ desviación típica de la distribución de probabilidad

Hacemos la prueba para una variable aleatoria continua $x:(-\infty,\infty)$.

La prueba sería parecida para el caso de una variable aleatoria discreta

Por definición $p(|x-\mu|\geq c) = \int_{|x-\mu|\geq c} f(x)\,dx$

$$\sigma^2 = \int_{\mathbb{R}} (x-\mu)^2 f(x)\,dx \geq \int_{|x-\mu|\geq c} (x-\mu)^2 f(x)\,dx \geq c^2 \int_{|x-\mu|\geq c} f(x)\,dx = c^2\, p(|x-\mu|\geq c) \Rightarrow$$

$$p(|x-\mu|\geq c)\leq \frac{\sigma^2}{c^2}$$

Si tomamos $c=2\sigma$ entonces obtenemos que $p(|x-\mu|\geq 2\sigma)\leq \frac{1}{4}$

Si tomamos $c=3\sigma$ entonces obtenemos que $p(|x-\mu|\geq 3\sigma)\leq \frac{1}{9}$

* Momentos de una distribución de probabilidad

$$\mu_k = E(x^k) = \begin{cases} \sum_i x_i^k f(x_i) & \text{(variable aleatoria discreta)} \\ \int x^k f(x)\,dx & \text{(variable aleatoria continua)} \end{cases}, \quad k = 1,2,\ldots \in \mathbb{N}$$

$\mu_1 =$ 1er momento = media de la distribución

$\mu_2 =$ 2º momento

$\vdots$

Podemos expresar la varianza en términos del 2º momento

$$V(x) = E((x-\mu)^2) = E((x-E(x))^2) = E(x^2 + E(x)^2 - 2xE(x)) = E(x^2) - 2E(x)^2 + E(x)^2 =$$

$$= E(x^2) - E(x)^2 = \mu_2 - \mu_1^2 = \sigma^2$$

►*A biased die has probabilities $p/2$, p, p, p, p, $2p$ of showing 1, 2, 3, 4, 5, 6 respectively. Find (i) the mean, (ii) the second moment and (iii) the variance of this probability distribution.*

Experimento aleatorio $\Rightarrow$ tirar un dado y ver el resultado

Espacio muestral 1, 2, 3, 4, 5, 6

Variable aleatoria x: $(1,2,3,4,5,6)$

Función de probabilidad

$$f(x_i) = \begin{cases} f(1) = p/2 \\ f(2) = p \\ f(3) = p \\ f(4) = p \\ f(5) = p \\ f(6) = 2p \end{cases}$$

$$\sum_i f(x_i) = \frac{p}{2} + 4p + 2p = \frac{p}{2} + 6p = \frac{13}{2}p = 1$$

$$p = \frac{2}{13}$$

- Media $\mu = E(x) = \sum_i x_i f(x_i) = \frac{p}{2} + 2p + 3p + 4p + 5p + 12p = \frac{p}{2} + 26p = \frac{53}{2}p = \frac{53}{2}\cdot\frac{2}{13} = \frac{53}{13}$

- 2º momento $\mu_2 = E(x^2) = \sum_i x_i^2 f(x_i) = \frac{p}{2} + 2^2p + 3^2p + 4^2p + 5^2p + 6^2 2p = \frac{p}{2} + 54p + 72p =$

$$= \frac{p}{2} + 126p = \frac{p}{2} + \frac{252}{2}p = \frac{253}{2}p = \frac{253}{2}\cdot\frac{2}{13} = \frac{253}{13}$$

- Varianza $\sigma^2 = \mu_2 - \mu^2 = \frac{253}{13} - \frac{53^2}{13^2} = \frac{253\times 13 - 2809}{13^2} = \frac{480}{169}$

$$\begin{array}{r} 53 \\ 53 \\ \hline 159 \\ 265 \\ \hline 2809 \end{array} \qquad \begin{array}{r} 253 \\ 13 \\ \hline 759 \\ 253 \\ \hline 3289 \\ -2809 \\ \hline 0480 \end{array} \qquad \begin{array}{r} 13 \\ 13 \\ \hline 39 \\ 13 \\ \hline 169 \end{array}$$

> Momentos centrales k-ésimos

$$\nu_k = E((x-\mu)^k) = E\left(\sum_{j=0}^{k} (-1)^j \binom{k}{j} x^{k-j}\mu^j\right) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \mu^j E(x^{k-j})$$

$$= \sum_{j=0}^{k} (-1)^j \binom{k}{j} \mu^j \mu_{k-j}$$

▸ Momentos centrales normalizados (adimensionales)

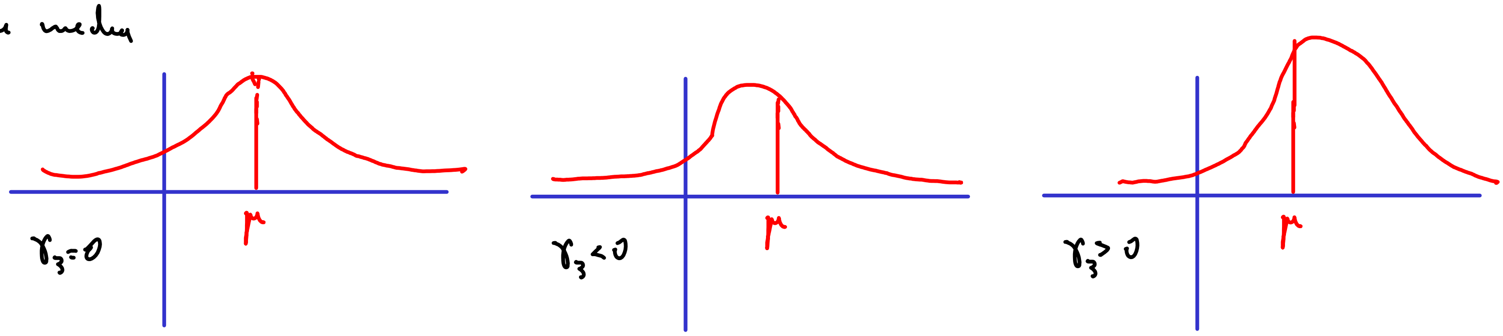
$$\gamma_k = \frac{\nu_k}{\nu_2^{k/2}} = \frac{\nu_k}{\sigma^k}$$

γ_3: coeficiente de asimetría

$\gamma_3 = 0$ distribución simétrica alrededor de la media

γ_4: curtosis o coeficiente de apuntamiento

Exceso de curtosis: $\gamma_4 - 3$



►*The PDF for a Gaussian distribution (see subsection 30.9.1) with mean μ and variance σ^2 is given by*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right].$$

Obtain an expression for the kth central moment of this distribution.

$$\nu_k = \int_{-\infty}^{\infty} (x-\mu)^k \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^k \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx =$$

$$= \left[\begin{array}{l} \frac{x-\mu}{\sigma\sqrt{2}} = y \\ dx = \sigma\sqrt{2}\,dy \end{array}\right] = \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (2)^{\frac{k}{2}} \sigma^k y^k e^{-y^2} dy = \frac{2^{\frac{k+1}{2}} \sigma^k}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^k e^{-y^2} dy$$

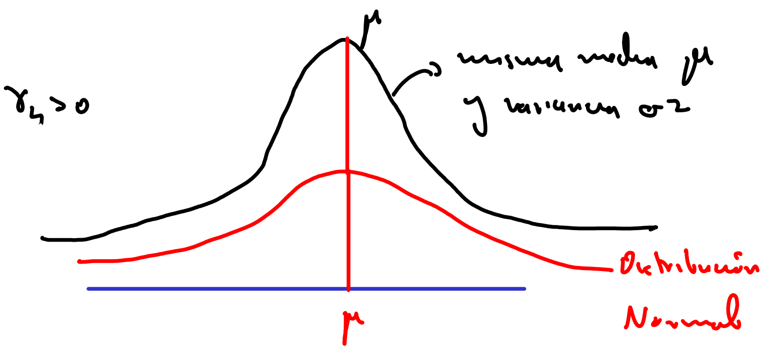
$$= \begin{cases} 0 & k \text{ impar} \\ \dfrac{2^{\frac{k+1}{2}} \sigma^k}{\sqrt{2\pi}} \sqrt{\pi}\, \dfrac{(k-1)(k-3)\cdots 3\cdot 1}{2^{k/2}} = \sigma^k (k-1)(k-3)\cdots 3\cdot 1 & (k \text{ par}) \end{cases}$$

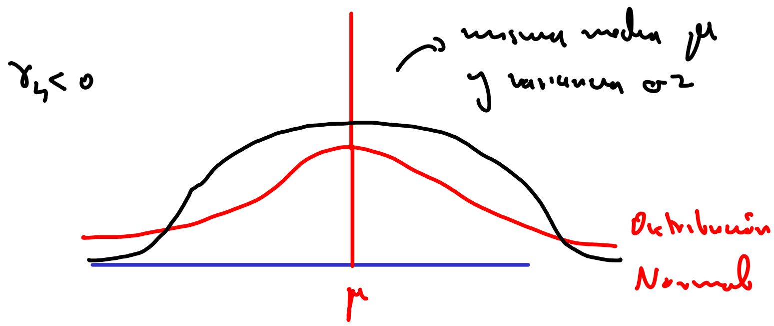
Por tanto $\gamma_3 = 0$, $\gamma_4 = \dfrac{\nu_4}{\sigma^4} = \dfrac{\sigma^4 \cdot 3 \cdot 1}{\sigma^4} = 3$

$$I(k)=\int_{-\infty}^{\infty} y^k e^{-y^2}\,dy = \int_{-\infty}^{\infty} y^{k-1}\frac{-2y}{-2} e^{-y^2}\,dy = \frac{-y^{k-1}}{2} e^{-y^2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} (k-1)\frac{y^{k-2}}{2} e^{-y^2}\,dy =$$

$$= \frac{(k-1)}{2} I(k-2) = \frac{(k-1)}{2}\frac{(k-3)}{2} I(k-4) = \dots \frac{(k-1)(k-3)\cdots(k-(k-1))}{2^{k/2}} I(0) =$$

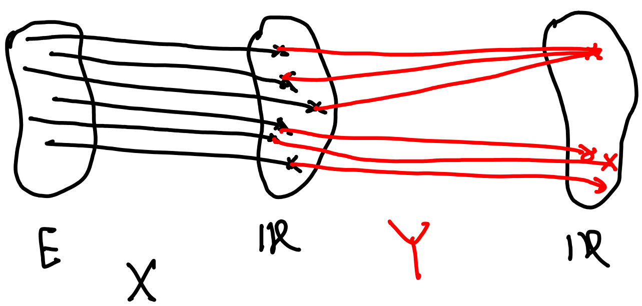
$$= \frac{(k-1)(k-3)\cdots 3\cdot 1}{2^{k/2}} \int_{-\infty}^{\infty} e^{-y^2}\,dy = \sqrt{\pi}\,\frac{(k-1)(k-3)\cdots 3\cdot 1}{2^{k/2}}$$





Funciones de variables aleatorias

Ya dijimos que una variable aleatoria X es una biyección entre el espacio muestral y un (sub)conjunto de los números reales. A veces nos puede interesar definir una nueva transformación de la variable aleatoria X en una nueva variable aleatoria Y



$$Y = Y(X)$$

¿Cómo se obtiene la distribución de probabilidad en términos de la nueva variable aleatoria Y?

- Función de probabilidad asociada a una transformación entre variables aleatorias discretas

Si la relación entre X e Y es biyectiva entonces se tiene la relación siguiente entre variables aleatorias

$$X: \{x_1 \dots x_n\} \longrightarrow Y: \left\{\begin{matrix} y(x_1) & \dots & y(x_n) \\ \shortparallel & & \shortparallel \\ y_1 & & y_n \end{matrix}\right\} \xrightarrow{\text{relación inversa}} X \left\{\begin{matrix} x(y_1) & \dots & x(y_n) \\ \shortparallel & & \shortparallel \\ x_1 & & x_n \end{matrix}\right\}$$

Utilizando las relaciones anteriores se obtiene la nueva función de probabilidad g en términos de la antigua f

$$f(x_i) = f(x(y_i)) = g(y_i)$$

Si la relación entre las variables aleatorias X e Y no es biyectiva entonces se tiene que:

$$X : \{x_1, x_2, \dots x_n\} \qquad Y : \left\{ \begin{array}{cccccc} \multicolumn{2}{c}{y_1} & \multicolumn{3}{c}{y_2} & y_3 \\ \nearrow & \nwarrow & \nearrow & \uparrow & \nwarrow & \uparrow \\ y(x_1) & y(x_2) & y(x_3) & y(x_4) & y(x_5) & y(x_n) \end{array} \right\}$$

$$g(y_1) = f(x_1) + f(x_2)$$

$$g(y_2) = f(x_3) + f(x_4) + f(x_5) \Rightarrow \text{En general} \quad g(y_i) \equiv \sum_{j \in S(y_i)} f(x_j)$$

$$g(y_3) = f(x_n)$$

$$S(y_i) = \{\text{índices } j \text{ tales que } y(x_j) = y_i$$

* Función densidad de probabilidad asociada a una transformación entre variables aleatorias continuas. De nuevo hay que distinguir entre el caso biyectivo y el caso no biyectivo.

$$x : I_x \subseteq \mathbb{R} \longrightarrow y : I_y \subseteq \mathbb{R}$$

$I_x \equiv$ dominio de $y = y(x)$

$I_y \equiv$ recorrido de $y = y(x)$

* Caso biyectivo

transformaciones directa e inversa $y = y(x)$, $x = x(y)$, $x(I_y) = I_x$, $y(I_x) = I_y$

$\Rightarrow$ Recorrido de $y(x)$ = dominio de $x(y)$

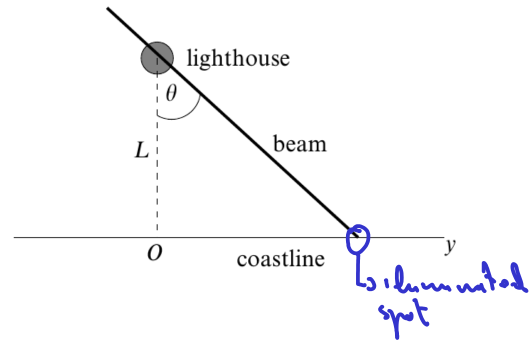
I_x e I_y definen los sucesos $x \in I_x$, $y \in I_y$. La biyectividad implica que $p(I_x) = p(I_y)$

Calculamos la probabilidad del suceso $x \in I_x$

$$p(I_x) = \int_{I_x} f(x)\,dx = \left[\begin{array}{l}\text{usar la fórmula} \\ \text{del cambio de variables} \\ \text{en una integral}\end{array}\right] = \int_{I_y} \overbrace{f(x(y))\,|x'(y)|}^{\text{nueva función densidad de probabilidad}}\,dy = p(I_y)$$

Ej. $I_x = [a, b] \Rightarrow \int_a^b f(x)\,dx = \int_A^B f(x(y))\,|x'(y)|\,dy \quad a = x(A),\ b = x(B)$

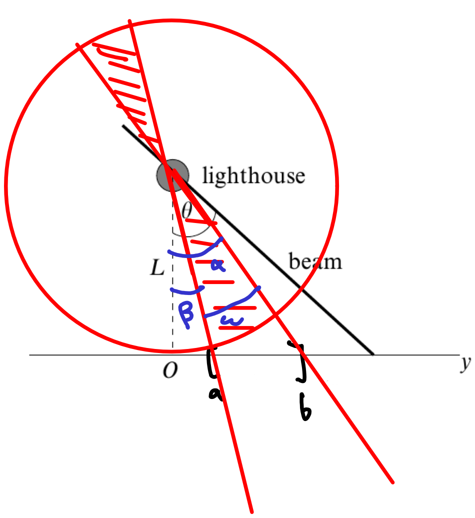
►*A lighthouse is situated at a distance L from a straight coastline, opposite a point O, and sends out a narrow continuous beam of light simultaneously in opposite directions. The beam rotates with constant angular velocity. If the random variable Y is the distance along the coastline, measured from O, of the spot that the light beam illuminates, find its probability density function.*



Experimento aleatorio = Situarnos en un punto de la línea de costa escogido al azar y comprobar si nos llega la luz del faro

Espacio muestral = Puntos de la línea de costa que reciben la luz del faro. Si tomamos un origen de coordenadas en O entonces la variable aleatoria se construye sin dificultad $y:(-\infty,\infty)$

Función densidad de probabilidad = Calculamos la probabilidad de recibir la luz del faro del suceso "situarse en la parte de la costa delimitada por las longitudes a, b" $A=\{a\leq y\leq b\}$



Como el haz de luz gira con velocidad angular constante, la probabilidad pedida será el área indicada del círculo de radio L dividida entre πL^2

$$\frac{b}{L}=\tan\alpha \quad \frac{a}{L}=\tan\beta \Rightarrow$$

$$\omega=\operatorname{ArcTan}\left(\frac{b}{L}\right)-\operatorname{ArcTan}\left(\frac{a}{L}\right)$$

Área de los sectores circulares: ωL^2

Área del sector circular $\Rightarrow \dfrac{\pi L^2}{2\pi}=\dfrac{A}{\omega} \Rightarrow A=\dfrac{\omega L^2}{2}$

Por tanto $p(A) = \frac{L^2}{\pi L^2}\left[\text{ArcTan}\left(\frac{b}{L}\right) - \text{ArcTan}\left(\frac{a}{L}\right)\right]$

Para calcular la función densidad de probabilidad suponemos que $b = a + h$, $a = y$ y $h \to 0$

$$\Rightarrow p(A) = \frac{1}{\pi}\left(\text{ArcTan}\left(\frac{y+h}{L}\right) - \text{ArcTan}\left(\frac{y}{L}\right)\right) \to \frac{1}{\pi}\left(\frac{\frac{1}{L}}{1+\left(\frac{y}{L}\right)^2}\right)h$$

Distribución de Cauchy
⇓

Así que la función densidad de probabilidad pedida es $f(y) = \frac{L^2}{\pi L (L^2+y^2)} = \frac{L}{\pi (L^2+y^2)}$

Comprobamos la normalización:

$$\int_{-\infty}^{\infty} \frac{L\,dy}{\pi(L^2+y^2)} = \begin{bmatrix} y = Lx \\ dy = L\,dx \end{bmatrix} = \frac{L}{\pi}\int_{-\infty}^{\infty}\frac{L\,dx}{L^2(1+x^2)} = \frac{1}{\pi}\left[\text{ArcTan}\, x\right]_{-\infty}^{\infty} =$$

$$= \frac{1}{\pi}\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = 1$$

* Es posible utilizar el ángulo ω como variable aleatoria en vez de la distancia al origen medida a lo largo de la línea de costa. La relación entre ambas es

$$\omega = \text{ArcTan}\left(\frac{b}{L}\right) - \text{ArcTan}\left(\frac{a}{L}\right) \Rightarrow a = 0 \quad b = y \Rightarrow \omega = \text{ArcTan}\left(\frac{y}{L}\right) \Rightarrow$$

$\Rightarrow y = L\tan\omega \Rightarrow$ relación biyectiva si $-\frac{\pi}{2} < \omega < \frac{\pi}{2}$. La función densidad de probabilidad en términos de la variable aleatoria ω es

$$g(\omega) = \frac{L}{\pi L^2(1+\tan^2\omega)} L \frac{1}{\sec^2\omega} =$$

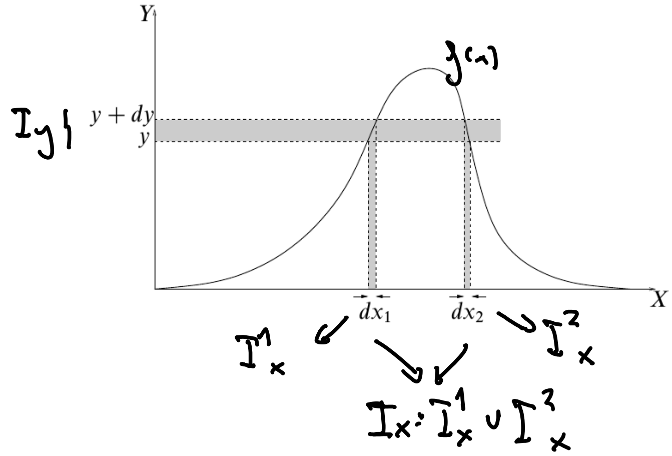
$$= \frac{\cos^4\omega}{\pi}$$

►*A lighthouse is situated at a distance L from a straight coastline, opposite a point O, and sends out a narrow continuous beam of light simultaneously in opposite directions. The beam rotates with constant angular velocity. If the random variable Y is the distance along the coastline, measured from O, of the spot that the light beam illuminates, find its probability density function.*

The situation is illustrated in figure 30.8. Since the light beam rotates at a constant angular velocity, θ is distributed uniformly between $-\pi/2$ and $\pi/2$, and so $f(\theta) = 1/\pi$. Now $y = L\tan\theta$, which possesses the single-valued inverse $\theta = \tan^{-1}(y/L)$, provided that θ lies between $-\pi/2$ and $\pi/2$. Since $dy/d\theta = L\sec^2\theta = L(1+\tan^2\theta) = L[1+(y/L)^2]$, from (30.58) we find

$$g(y) = \frac{1}{\pi}\left|\frac{d\theta}{dy}\right| = \frac{1}{\pi L[1+(y/L)^2]} \qquad \text{for } -\infty < y < \infty.$$

* La relación entre las variables aleatorias $x \longleftrightarrow y$ no es inyectiva



La función $y=y(x)$ no es inyectiva $\Rightarrow$ su restricción a I_x^1 o I_x^2 sí es biyectiva

$\exists\ x^1 = x^1(y)$, $x^2 = x^2(y)$ tales que

$y(I_x^1) = I_y \Rightarrow x^1(I_y) = I_x^1$

$y(I_x^2) = I_y \Rightarrow x^2(I_y) = I_x^2$

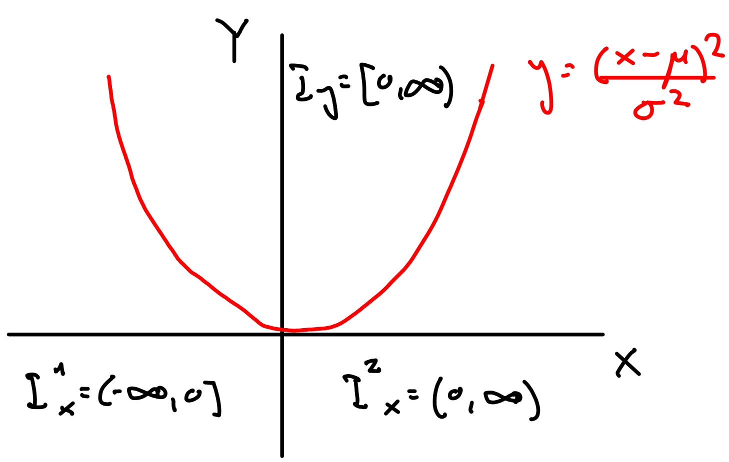
$x^1(y)$, $x^2(y)$ se denominan ramas

$$p(I_y) = p(I_x^1) + p(I_x^2) = \int_{I_x^1} f(x)\,dx + \int_{I_x^2} f(x)\,dx = \left[\begin{array}{l}\text{Usar fórmula del} \\ \text{cambio de variables} \\ \text{en cada trozo}\end{array}\right] =$$

$$= \int_{I_y} f(x^1(y))\,|x^1(y)'|\,dy + \int_{I_y} f(x^2(y))\,|x^2(y)'|\,dy \Rightarrow$$

$$= \int_{I_y} \Big(\underbrace{f(x^1(y))\,|x^1(y)'| + f(x^2(y))\,|x^2(y)'|}_{\text{nueva función densidad de probabilidad}} \Big)\,dy \Rightarrow g(y) = f(x^1)\,|x^1(y)'| + f(x^2(y))\,|x^2(y)'|$$

► *The random variable X is Gaussian distributed (see subsection 30.9.1) with mean μ and variance σ^2. Find the PDF of the new variable $Y = (X-\mu)^2/\sigma^2$.*



Distribución gaussiana:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right]$$

$y = y(x)$ es biyectiva en cada uno de los intervalos I_x^1, I_x^2

$$y = \frac{(x-\mu)^2}{\sigma^2} \Rightarrow \pm\sigma\sqrt{y} + \mu = x = \begin{cases} x^2(y) = \sigma\sqrt{y} + \mu \Rightarrow x^2(I_y) = I_x^2 \\ x^1(y) = -\sigma\sqrt{y} + \mu \Rightarrow x^1(I_y) = I_x^1 \end{cases}$$

Así que:

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{y}{2}}\left(|x^2(y)'| + |x^1(y)'|\right) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{y}{2}}\left(\frac{\sigma}{2\sqrt{y}} + \frac{\sigma}{2\sqrt{y}}\right) = \frac{1}{\sqrt{2\pi y}}\, e^{-\frac{y}{2}} =$$

$$= \frac{1}{2\left(\frac{\pi y}{2}\right)^{\frac{1}{2}}}\, e^{-\frac{y}{2}} \Rightarrow \text{distribución gamma } \gamma\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

* Las definiciones analizadas se extienden sin dificultad al caso de variables aleatorias multidimensionales. P. ej. sea (x, y) una variable aleatoria multidimensional con una distribución de probabilidad dada y definimos una transformación $z = z(x, y)$ ¿Cuál será la distribución de probabilidad asociada a la nueva variable aleatoria z?

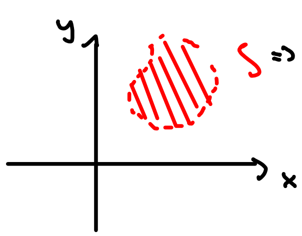
* x, y son variables aleatorias discretas ⇒

$x: \{x_1, \dots x_n\}$, $y: \{y_1, \dots y_m\}$, función de probabilidad $f(x_i, y_j)$

Nueva variable aleatoria $z: \{z_{11}, \dots, z_{nm}\}$, $z_{11} = z(x_1, y_1)$, $z_{12} = z(x_1, y_2)$, $z_{1m} = z(x_1, y_m)$, $\dots$ $z_{nm} = z(x_n, y_m)$

$$p(z) = \sum_{i,j} f(x_i, y_j) \quad \text{donde } (x_j, y_j) \text{ son tales que } z = z(x_i, y_j)$$

* x, y son variables aleatorias continuas ⇒ $z = z(x, y)$ ⇒ función densidad de probabilidad $f(x, y)$



S ⇒ la región S del plano x-y define un suceso ⇒ $p(S) = \int_S f(x,y)\,dx\,dy$

$I \equiv z(S)$ si $g(z)$ es la función densidad de probabilidad de la nueva variable aleatoria z entonces

$z: \mathbb{R}^2 \to \mathbb{R}$

$p(S) = \int_I g(z)\,dz$ ⇒ usar la relación

$\int_S f(x,y)\,dx\,dy = \int_I g(z)\,dz$ para hallar $g(z)$

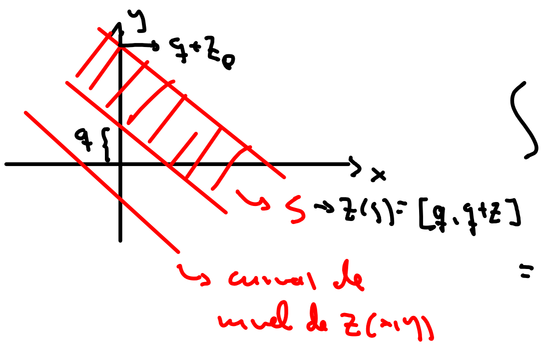
Lo mejor es verlo con un ejemplo.

►*Suppose X and Y are independent continuous random variables in the range $-\infty$ to ∞, with PDFs $g(x)$ and $h(y)$ respectively. Obtain expressions for the PDFs of $Z = X + Y$ and $W = XY$.*

Si X e Y son variables aleatorias continuas e independientes entonces la función densidad de probabilidad es $f(x,y) = g(x)h(y)$ $\quad x:(-\infty,\infty)$, $y:(-\infty,\infty)$

$$\int_I p(z)\,dz = \int_S f(x,y)\,dx\,dy \qquad z(S) = I$$

$z = x+y$



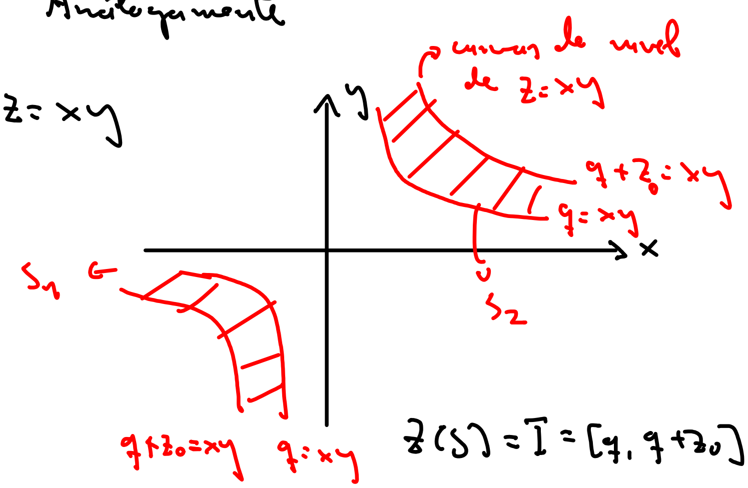
$$\int_S f(x,y)\,dx\,dy = \int_{-\infty}^{\infty} dx \int_{q-x}^{q+z_0-x} dy\, g(x)h(y) =$$

$$= \int_{-\infty}^{\infty} dx\, g(x) \int_{q-x}^{q+z_0-x} dy\, h(y) = \int_q^{q+z_0} p(s)\,ds = F(z_0)$$

$$\Rightarrow F'(z_0) = p(q+z_0) = \int_{-\infty}^{\infty} dx\, g(x)\, h(q+z_0-x) \Rightarrow p(z_0) = \int_{-\infty}^{\infty} dx\, g(x)\, h(z_0-x)$$

Análogamente

$z = xy$



$S = S_1 \cup S_2$

$$\int_S f(x,y)\,dx\,dy = \int_{-\infty}^{\infty} dx \int_{\frac{q}{x}}^{\frac{q+z_0}{x}} dy\, f(x)g(y) =$$

$$= \int_{-\infty}^{\infty} dx\, f(x) \int_{\frac{q}{x}}^{\frac{q+z_0}{x}} dy\, g(y) = \int_q^{q+z_0} p(s)\,ds = F(z_0)$$

$z(S) = I = [q, q+z_0]$

$$F'(z_0) = p(q+z_0) = \int_{-\infty}^{\infty} dx\, f(x)\frac{1}{x}\, g\left(\frac{q+z_0}{x}\right) \Rightarrow p(z) = \int_{-\infty}^{\infty} dx\, f(x)\frac{1}{x}\, g\left(\frac{z}{x}\right)$$

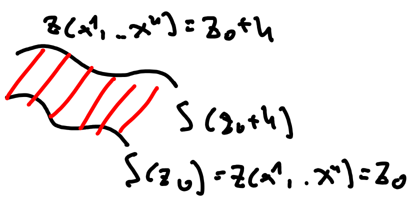
* Este método se generaliza sin dificultad al caso con múltiples variables aleatorias $x_1, \dots x_n$ $\quad z = z(x_1, \dots x_n)$

- Caso discreto $\quad p(z_k) = \sum_{j_1 \dots j_n} z_k(x_{j_1}, \dots, x_{j_n})$ donde $j_1 \dots j_n$ son tales que $z_k = z_k(x_{j_1}, \dots, x_{j_n})$

- Caso continuo $\quad \int_I g(z)\, dz = \int_S f(x^1, \dots x^n)\, dx^1 \dots dx^n$ donde $z(S) = I$, $S \subset \mathbb{R}^n$

Consideramos una región $S(z_0+h) \subset \mathbb{R}^n$ tal que $z[S(z_0+h)] = [z_0, z_0+h]$, es decir:

$$S(z_0+h) = \{ (x^1, \dots, x^n) \in \mathbb{R}^n : z_0 \leq z(x^1, \dots, x^n) \leq z_0 + h \}$$



$\Rightarrow S(z_0+h)$ es la región comprendida entre 2 superficies de nivel $z(x^1, \dots x^n) = z_0$ y $z(x^1, \dots x^n) = z_0 + h$

Entonces se tiene que

$$\lim_{h \to 0} \frac{1}{h} \Big[\int_{S(z_0+h)} f(x^1, \dots x^n)\, dx^1 \dots dx^n - \underbrace{\int_{S(z_0)} f(x^1, \dots, x^n)\, dx^1, \dots dx^n}_{0} \Big] = \lim_{h \to 0} \int_{z_0}^{z_0+h} g(z)\, dz =$$

$= g(z_0)$ $\qquad$ Pero por otro lado

$$\lim_{h \to 0} \frac{1}{h} \int_{S(z_0+h)} f(x^1, \dots x^n)\, dx^1 \dots dx^n = \int_{\mathbb{R}^n} f(x^1, \dots x^n)\, \delta(z(x^1, \dots, x^n) - z_0)\, dx^1 dx^2 \dots dx^n$$

Así que $\boxed{g(z_0) = \int_{\mathbb{R}^n} f(x^1, \dots x^n)\, \delta(z(x^1, \dots, x^n) - z_0)\, dx^1 \dots dx^n}$

Recordemos la siguiente representación de la delta $\delta(z(x^1, \dots x^n) - z_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(z(x^1, \dots x^n) - z_0)}\, dk$

También se puede hacer sin delta de Dirac:

$$\int_{\mathbb{R}^n} f(x^1,..x^n)\,\delta(z(x^1,..,x^n)-z_0) = \int_{S(z_0)} f(x^1,..,x^n)\,d\mu_{S(z_0)}$$

$d\mu_{S(z_0)} \equiv$ medida inducida en $S(z_0)$ por $d^nV = dx^1..dx^n$ que es la medida utilizada para la integral estándard (Riemann o Lebesgue) en $\mathbb{R}^n$

La relación con la delta de Dirac viene dada por la propiedad de la delta:

$$\langle \delta(z(x^1,...,x^n)-z_0),\, \psi(x^1,..x^n)\rangle = \int_{S(z_0)} \psi\, d\mu_{S(z_0)}$$

►*A general one-dimensional random walk consists of n independent steps, each of which can be of a different length and in either direction along the x-axis. If g(x) is the PDF for the (positive or negative) displacement X along the x-axis achieved in a single step, obtain an expression for the PDF of the total displacement S after n steps.*

- Experimento aleatorio ⇒ Medir la posición de una partícula que se mueve aleatoriamente a lo largo del eje x
- Espacio muestral $\mathbb{R}: (-\infty, \infty)$
- Variable aleatoria $x: (-\infty, \infty)$
- Función densidad de probabilidad $g(x)$

⇒ Ahora repetimos el experimento aleatorio n veces

- Espacio muestral $\mathbb{R}\times\mathbb{R}\times ... \times\mathbb{R} = \mathbb{R}^n$
- Variable aleatoria $(x^1,..,x^n): \mathbb{R}^n$
- Función densidad de probabilidad ⇒ los n experimentos aleatorios son independientes así que $f(x^1,..x^n) = g(x^1)\,g(x^2)\cdots g(x^n)$

- Finalmente nos preguntamos por la función densidad de probabilidad del "desplazamiento total" que denotaremos por S. Dado que $S = x^1 + x^2 + \dots + x^n$ se tiene que necesitamos calcular la función densidad de probabilidad $p(s)$ que resulta de hacer la transformación de variables aleatorias $S = x^1 + \dots + x^n$. Aplicamos la fórmula

$$p(s_0) = \int_{\mathbb{R}^n} f(x^1, \dots, x^n)\, \delta(S(x^1, \dots, x^n) - s_0)\, dx^1 \dots dx^n$$

$$= \int_{\mathbb{R}^n} f(x^1, \dots, x^n)\, \delta(x^1 + x^2 + \dots + x^n - s_0)\, dx^1 \dots dx^n =$$

$$\int_{\mathbb{R}^n} f(x^1 \dots x^n)\, dx^1 \dots dx^n \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x^1 + x^2 + \dots + x^n - s_0)}\, dk$$

$$= \int_{-\infty}^{\infty} dk \frac{1}{2\pi} e^{-iks_0} \int_{-\infty}^{\infty} dx^1 e^{ikx^1} \int_{-\infty}^{\infty} dx^2 e^{ikx^2} \dots \int_{-\infty}^{\infty} dx^n e^{ikx^n} \underbrace{f(x^1, \dots, x^n)}_{g(x^1) \dots g(x^n)}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-iks_0} \left[\int_{-\infty}^{\infty} dx^1 e^{ikx^1} g(x^1)\right] \left[\int_{-\infty}^{\infty} dx^2 e^{ikx^2} g(x^2)\right] \dots \left[\int_{-\infty}^{\infty} dx^n e^{ikx^n} g(x^n)\right]$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-iks_0} \underbrace{\left[\int_{-\infty}^{\infty} dx\, e^{ikx} g(x)\right]}_{\substack{C_x(k) \equiv \text{función característica} \\ \text{de la variable} \\ \text{aleatoria } x}}{}^{n} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-iks_0} \left[C(k)\right]^n = p(s_0)$$

Funciones generatrices

* Función generatriz de probabilidad

Condiciones $\Rightarrow$ variable aleatoria $x: (1, 2, \ldots k)$, $k \in \mathbb{N}$
función de probabilidad $f(n)$

Suponemos sin pérdida de generalidad que el recorrido de la variable aleatoria es $\mathbb{N}$
En estas condiciones, la función generatriz de probabilidad se define así:

$$\phi_x(t) = \sum_{n=1}^{\infty} f(n) t^n = E(t^n)$$

Propiedades elementales: $\phi_x(1) = \sum_{n=1}^{\infty} f(x_n) = 1$

$$\phi_x'(t) = \sum_{n=1}^{\infty} n f(x_n) t^{n-1} \Rightarrow \phi_x'(1) = \sum_{n=1}^{\infty} n f(x_n) = E(n) = \mu \qquad \boxed{\mu = \phi_x'(1)}$$

$$\phi_x''(t) = \sum_{n=1}^{\infty} n(n-1) f(x_n) t^{n-2} \Rightarrow \phi_x''(1) = \sum_{n=1}^{\infty} n^2 f(x_n) - \sum_{n=1}^{\infty} n f(x_n) = E(n^2) - E(n) = E(n^2) - \mu$$

$\sigma^2 =$ Varianza.

$$\Rightarrow E(n^2) = \phi_x''(1) + \mu = \phi_x''(1) + \phi_x'(1)$$

$\sigma =$ desviación típica.

Podemos utilizar la función generatriz para calcular σ^2

$$\sigma^2 = E\left[(n-\mu)^2\right] = E(n^2 + \mu^2 - 2n\mu) = E(n^2) - 2\mu E(n) + \mu^2 = E(n^2) - \mu^2$$

$$= \phi_x''(1) + \phi_x'(1) - \phi_x'(1)^2 \Rightarrow \boxed{\sigma^2 = \phi_x''(1) + \phi_x'(1) - \phi_x'(1)}$$

> Función generatriz de momentos

$$M_x(t) = \begin{cases} \sum_i e^{t x_i} f(x_i) & \text{(Variable aleatoria discreta)} \\ \int e^{tx} f(x)\,dx & \text{(Variable aleatoria continua)} \end{cases}$$

* Propiedades elementales:

$$M_X(t) = E(e^{tX})$$

$$M_X(0) = 1 \quad , \quad M_X'(0) = E(X) = \mu$$

$$M_X^{(n)}(0) = E(X^n)$$

* Si la función generatriz de probabilidad está definida entonces

$$\phi_X(e^t) = \sum_{n=1}^{\infty} f(n) e^{nt} = M_X(t)$$

* Relación con la función característica: $M_X(it) = C_X(t) \quad , \quad C_X(-it) = M_X(t)$

* $M_{aX+b}(t) = E(e^{t(aX+b)}) = E(e^{atX} e^{bt}) = e^{bt} E(e^{atX}) = e^{bt} M_{aX}(t)$

Esto se puede utilizar para obtener los momentos centrales de una distribución:

$$M_{X-\mu}(t) = e^{-\mu t} M_X(t) \Rightarrow E((X-\mu)^n) = M_{X-\mu}^{(n)}(0) = \frac{d^n}{dt^n}\left[e^{-\mu t} M_X(t)\right]\Big|_{t=0}$$

* Cálculo de la desviación típica/varianza

$$E(X^2) = M_X''(0) \Rightarrow \sigma^2 = E(X^2) - E(X)^2 = M_X''(0) - M_X'(0)^2$$

* Suma de variables aleatorias: si tenemos una suma de variables aleatorias independientes $X = X_1 + \dots + X_n$. Entonces:

$$M_X(t) = E(e^{tX}) = E(e^{tX_1} \dots e^{tX_n}) = E(e^{tX_1}) \dots E(e^{tX_n}) =$$

$$= \prod_{i=1}^{n} M_{X_i}(t)$$

y si se trata de una combinación lineal de variables aleatorias independientes $X = \sum_i c_i X_i$

$$M_X(t) = E(e^{\sum_i c_i X_i t}) = E\left(\prod_{i=1}^{n} e^{c_i X_i t}\right) = \prod_{i=1}^{n} M_{X_i}(c_i t)$$

Ejemplos de distribuciones de probabilidad

- Distribución binomial

- Experimento aleatorio ⇒ experimento aleatorio compuesto por n experimentos aleatorios sucesivos simples e independientes. Cada experimento aleatorio simple tiene 2 posibles resultados A y $\bar{A}$ (A = éxito, $\bar{A}$ = fracaso).

Experimento aleatorio simple 1: A / $\bar{A}$ Experimento aleatorio simple 2: A / $\bar{A}$ . . . . . Experimento aleatorio simple n: A / $\bar{A}$

- Espacio muestral ⇒ $\{\underbrace{A, \dots A}_{n}\}$, $\{\underbrace{A A}_{n-1}\ \bar{A}\}$, $\{\underbrace{A \dots A}_{n-2}\ \bar{A}\bar{A}\}$, $\{\overbrace{AA \cdot A}^{x}\ \overbrace{\bar{A} \cdot \bar{A}}^{n-x}\}$, $\{A\underbrace{\bar{A} \cdots \bar{A}}_{n-1}\}$, $\{\underbrace{\bar{A} \cdots \bar{A}}_{n}\}$

Lo dividimos en bloques:

$\{\underbrace{A, \dots A}_{n}\} \leftarrow PR^n_n$ (n éxitos)

$\{A A \cdot \bar{A} A\} \leftarrow PR^n_{n-1,1} \rightarrow \{\bar{A}\underbrace{A \dots A}_{n-1}\}$ ($n-1$ éxitos)

$\{\bar{A} A \cdot\cdot A \bar{A}\} \leftarrow PR^n_{n-2,2} \rightarrow \{\bar{A}\bar{A} A \cdots A\}$ ($n-2$ éxitos)

$PR^n_{x,n-x} \rightarrow$ x éxitos

$PR^n_{1,n-1} \rightarrow$ 1 éxito

$PR^n_n \rightarrow$ 0 éxitos.

En cada bloque nos importa el orden porque nos importa el orden de los sucesivos experimentos aleatorios

$$\text{Nº total de elementos del espacio muestral} = \sum_{x=0}^{n} PR^n_{x,n-x} = \sum_{x=0}^{n} \frac{n!}{x!(n-x)!} = \sum_{x=0}^{n} \binom{n}{x} = 2^n = VR^2_n$$

- Variable aleatoria = nº de éxitos (o nº de fracasos) $X: (1, 2, 3, \dots, n)$

- Función de probabilidad: Si r es la probabilidad de éxito A en un único experimento aleatorio entonces $p(\bar{A}) = 1 - p(A) = 1 - r$

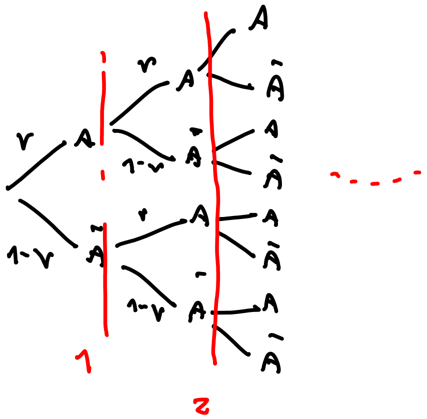
Probabilidad de un elemento del bloque $\rightarrow$ $p\{\underbrace{A A}_{x} \cdot \underbrace{\bar{A} \cdot \bar{A}}_{n-x}\} = p(A)^x\, p(\bar{A})^{n-x}$ ⇒ $\boxed{p(x) = \binom{n}{x} r^x (1-r)^{n-x}}$

↳ Distribución binomial

$X \sim B(n, r)$

diagrama de probabilidades

↓



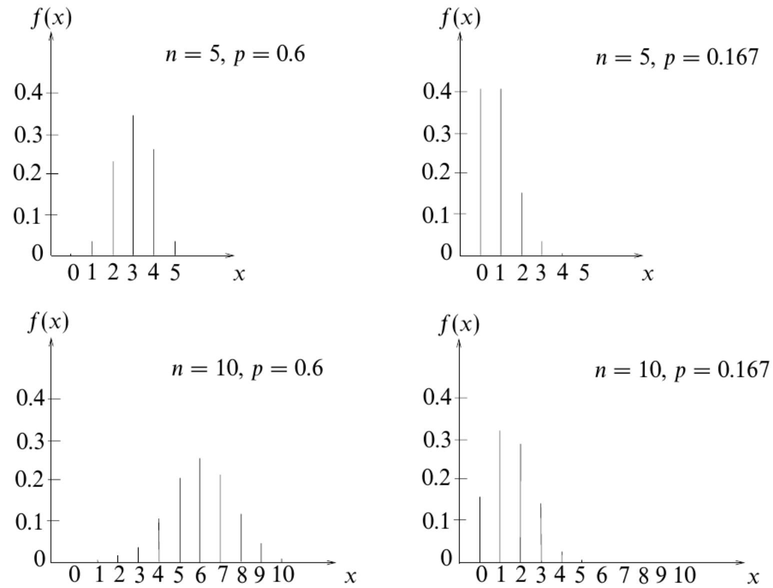


Figure 30.11 Some typical binomial distributions with various combinations of parameters n and p.

* función generatriz de momentos

Fórmula del binomio ↓

$$M_x(t) = \sum_{x=0}^{n} e^{xt} p(x) = \sum_{x=0}^{n} e^{xt} \binom{n}{x} r^x (1-r)^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (re^t)^x (1-r)^{n-x} =$$

$$= (re^t + 1 - r)^n \Rightarrow$$

* Media y varianza (desviación típica)

$$M_x'(t) = n(re^t + 1 - r)^{n-1} re^t \qquad \mu = M_x'(0) = nr$$

$$M_x''(t) = n(n-1)(re^t + 1 - r) r^2 e^{2t} + n(re^t + 1 - r)^{n-1} re^t \Rightarrow M_x''(0) = n(n-1)r^2 + nr$$

$$\sigma^2 = n(n-1)r^2 + nr - n^2 r^2 = nr - nr^2 = nr(1-r)$$

►*If a single six-sided die is rolled five times, what is the probability that a six is thrown exactly three times?*

Experimento aleatorio simple: lanzar un dado de 6 caras $\Rightarrow$ nº de veces que se repite el experimento aleatorio elemental: 5

éxito $\Rightarrow$ sale seis $\Rightarrow r = \frac{1}{6}$

$n = 5$

$$p(x=3) = \binom{n}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{n-3} = \binom{5}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^2 = \frac{5!}{3!\,2!} \frac{5^2}{6^5} = \frac{5^3 \cdot 4}{2 \cdot 6^5} = \frac{250}{7776} \simeq 0.0322$$

- Distribución geométrica

Experimento aleatorio => experimento aleatorio compuesto por infinitos experimentos aleatorios sucesivos simples e independientes. Cada experimento aleatorio simple tiene 2 posibles resultados A y $\bar{A}$ (A = éxito, $\bar{A}$ = fracaso).

Experimento aleatorio simple 1 — A / $\bar{A}$; Experimento aleatorio simple 2 — A / $\bar{A}$ .. Experimento aleatorio simple n — A / $\bar{A}$ . . . .

El experimento aleatorio se repite hasta obtener un éxito.

			Nº de veces que se repite el experimento aleatorio
	A	-> No hay fracasos	1
- Espacio muestral =	$\bar{A}A$	-> El segundo resultado es éxito	2
El espacio muestral es infinito	$\bar{A}\bar{A}A$	-> El tercer resultado es éxito	3
	$\bar{A}\bar{A}\bar{A}A$	-> el 4º resultado es éxito	4
	$\bar{A}\bar{A}\bar{A}\bar{A}A$	-> el 5º resultado es éxito	5
	$\vdots$		

- Variable aleatoria = nº de veces que se repite el experimento aleatorio hasta obtener un éxito

$X: (1, 2, 3, 4, \ldots)$ El recorrido de la variable aleatoria es el conjunto de los números naturales.

- función de probabilidad $= p(x) = p(\underbrace{\bar{A}\cap\bar{A}\cap\ldots\bar{A}}_{x-1}\cap A) = p(\bar{A})^{x-1}\cdot p(A) = (1-r)^{x-1}\cdot r$

- función generatriz de momentos

$$M_x(t) = E(e^{xt}) = \sum_{x=1}^{\infty} e^{xt}(1-r)^{x-1}r = \sum_{x=1}^{\infty}\left[e^t(1-r)\right]^x \frac{r}{1-r} = \frac{r\,e^t(1-r)}{(1-e^t(1-r))(1-r)}$$

$$= \frac{r\,e^t}{1-e^t(1-r)}$$

$$M_x(t) = \frac{re^t}{1-e^t(1-r)} \qquad M'_x(t) = \frac{re^t(1-e^t(1-r)) + re^t(1-r)e^t}{(1-e^t(1-r))^2} =$$

$$= \frac{re^t - r(1-r)e^{2t} + r(1-r)e^{2t}}{(1-(1-r)e^t)^2} = \frac{re^t}{(1-(1-r)e^t)^2}$$

$$M''_x(t) = \frac{re^t(1-(1-r)e^t)^2 + re^t\,2(1-(1-r)e^t)(1-r)e^t}{(1-(1-r)e^t)^4} =$$

$$= \frac{[re^t - (1-r)re^{2t} + 2r(1-r)e^{2t}](1-(1-r)e^t)}{(1-(1-r)e^t)^4} = \frac{re^t[1-(1-r)^2e^{2t}]}{(1-(1-r)e^t)^4}$$

De aquí calculamos la media y varianza/desviación típica

$$\mu = M'_x(0) = \frac{1}{r}$$

$$\sigma^2 = M''_x(0) - M'_x(0)^2 = \frac{r[1-(1-r)^2]}{(r^4)} - \frac{1}{r^2} = \frac{r\,r(2-r)}{r^4} - \frac{1}{r^2} = \frac{1-r}{r^2}$$

$$\lim_{n\to\infty} (e^t(1-r))^{n+1} = \begin{cases} 0 & e^t(1+r) < 1 \Rightarrow e^t < \frac{1}{1+r} \Rightarrow t < -\log[1+r] \\ \infty & e^t(1+r) > 1 \end{cases}$$

Recordemos la fórmula de la suma de una progresión geométrica

$$\{rR, r^2R, \ldots, r^nR\}$$

$$S = rR + r^2R + \ldots + r^nR \Rightarrow \quad rS = r^2R + \ldots r^nR + r^{n+1}R \quad = S + r^{n+1}R - rR$$

$$rS = S + R(r^{n+1} - r) \Rightarrow S(r-1) = R(r^{n+1}-r) \Rightarrow S = \frac{R(r^{n+1}-r)}{r-1}$$

$$\lim_{n\to\infty} S = -\frac{rR}{r-1} = \frac{Rr}{1-r}$$

* Distribución binomial negativa

* Experimento aleatorio ⇒ experimento aleatorio compuesto por infinitos experimentos aleatorios sucesivos simples e independientes. Cada experimento aleatorio simple tiene 2 posibles resultados A y $\bar{A}$ (A = éxito, $\bar{A}$ = fracaso).

$$\underset{1}{\text{Experimento aleatorio simple}} \begin{cases} A \\ \bar{A} \end{cases} \quad \underset{2}{\text{Experimento aleatorio simple}} \begin{cases} A \\ \bar{A} \end{cases} \quad \cdots \quad \text{Experimento aleatorio simple} \begin{cases} A \\ \bar{A} \end{cases} \quad \cdots$$

Además se supone que el resultado del último experimento aleatorio es un éxito y que ha habido k éxitos

* Espacio muestral

$\underbrace{A \cdots A}_{k} A \cdots$	$\overbrace{\bar{A} \cdots A}^{k+1} A \cdots$	$\overbrace{\bar{A}\bar{A} A \cdots A}^{k+2} A \cdots$	$\cdots$	$\overbrace{\bar{A}\bar{A}\bar{A} \cdots \quad A}^{k+x} \cdots$
	$A\bar{A} \cdots A$	$\bar{A}A\bar{A} \cdots A$		$\vdots$
	$\vdots$	$A\bar{A}\bar{A} \cdots A$		$PR^{k+x-1}_{k,x}$
	$A \quad \bar{A}A$	$\vdots$		$\downarrow$
0 fracasos	1 fracaso	2 fracasos		x fracasos

* Variable aleatoria = nº de fracasos

* Función de probabilidad $p(x) = \underbrace{p(\bar{A}) \cdots p(\bar{A})}_{x} \cdot \underbrace{p(A) \cdots p(A)}_{k-1} \cdot p(A) \cdot PR^{x+k-1}_{x,k-1}$

$$= (1-r)^x \, r^{k-1} \, r \binom{x+k-1}{x} = (1-r)^x r^k \binom{x+k-1}{x}$$

* Función generatriz de momentos

$$M_x(t) = E(e^{xt}) = \sum_{x=0}^{\infty} e^{xt} (1-r)^x r^k \binom{x+k-1}{x} = r^k \sum_{x=0}^{\infty} (e^t(1-r))^x \binom{x+k-1}{x}$$

Ahora necesitamos recordar la fórmula siguiente:

$$(a-b)^{-n} = \sum_{h=0}^{\infty} \binom{n-1+h}{h} \left(\frac{b}{a^n}\right)^h \qquad |a| > |b| \qquad \Rightarrow \text{utilizando la fórmula} \Rightarrow$$

$$\Rightarrow M_x(t) = \frac{r^k}{(1-e^t(1-r))^k} = \left[\frac{r}{(1-e^t(1-r))}\right]^k$$

$$M_x'(t) = r^k(-k)\left[1-e^t(1-r)\right]^{-k-1}(-e^t(1-r)) = r^k k(1-r)e^t\left[1-e^t(1-r)\right]^{-1-k}$$

$$M_x''(t) = r^k k(1-r) e^t \left[\left[1-e^t(1-r)\right]^{-1-k} + (1+k)\left[1-e^t(1-r)\right]^{-2-k} e^t(1-r) \right]$$

$$\mu = M_x'(0) = r^k k(1-r)[r]^{-1-k} = k\frac{(1-r)}{r}$$

$$\sigma^2 = M_x''(0) - M_x'(0)^2 = r^k k(1-r)\left[\frac{1}{r^{1+k}} + (1+k)\frac{(1-r)}{r^{2+k}}\right] - k^2\frac{(1-r)^2}{r^2}$$

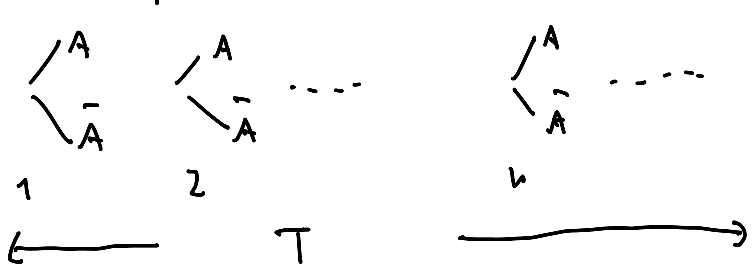
$$= \frac{k(1-r)}{r} + \frac{k(1-r)^2}{r^2}(1+k) - k^2\frac{(1-r)^2}{r^2} = \frac{k(1-r)}{r} + k\frac{(1-r)^2}{r^2} = \frac{k(1-r)}{r}\left[1 + k\frac{(1-r)}{r}\right]$$

$$= \frac{k(1-r)}{r}\,\frac{r+k-kr}{r} = \frac{k(1-r)}{r^2}(k + r(1-k))$$

hay algo mal → debe ser 1 $\quad \sigma^2 = \frac{k(1-r)}{r^2}$

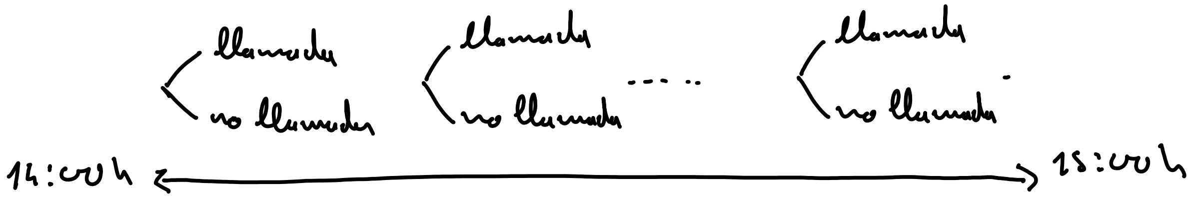
Distribución de Poisson

• Experimento aleatorio: sucesión infinita de experimentos aleatorios elementales. Cada uno de los experimentos aleatorios elementales admite dos posibles resultados: éxito o fracaso



Los experimentos aleatorios elementales tienen lugar dentro de un intervalo de tiempo que es finito.
El número medio de éxitos en dicho intervalo de tiempo es un parámetro conocido

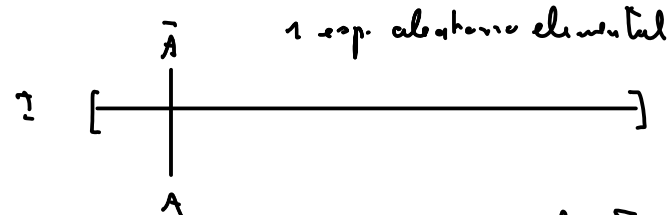
* Ejemplos: Véanse cuantas llamadas recibe una teleoperadora de 2 a 3 de la tarde.

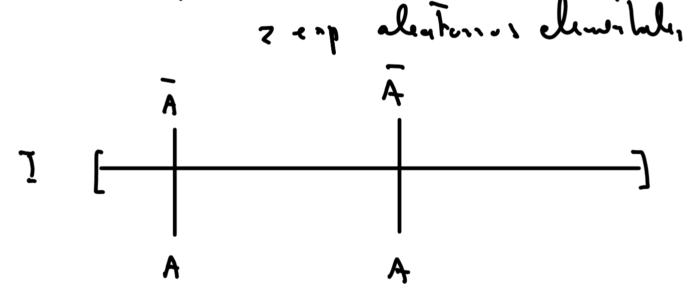


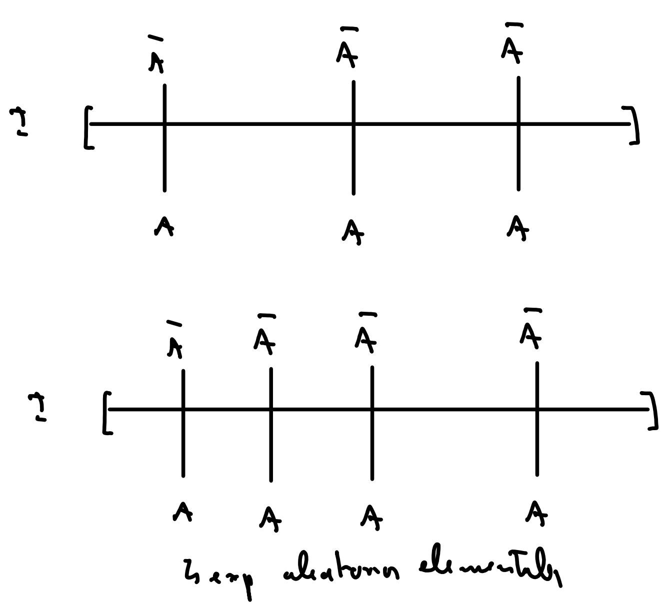
Los experimentos aleatorios elementales pueden suceder en cualquier instante entre las 14:00h y las 15:00h

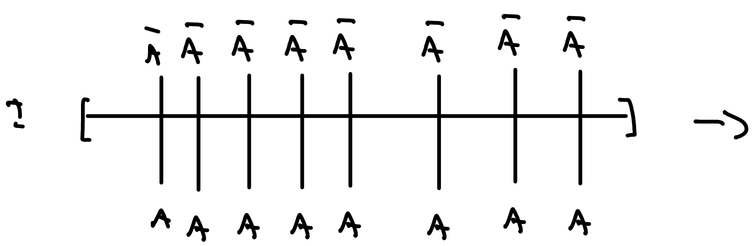
* Espacio muestral: el espacio muestral es infinito

$I \subseteq \mathbb{R}$ intervalo (intervalo temporal)









8 experimentos aleatorios elementales

El espacio muestral sería el límite cuando el número de experimentos aleatorios elementales cubre el intervalo I =

$$E = \bigcup_{a \in I} \{a, A, \bar{A}\},$$ a = hora a la que se hace el experimento aleatorio.

Ejemplo de suceso con 2 éxitos

$$(a_1, A), (a_2, A), \bigcup_{a \in I \setminus \{a_1 \cup a_2\}} (a, \bar{A})$$

* Variable aleatoria $x =$ nº de éxitos $x: (0, 1, 2, \ldots)$ en el intervalo de tiempo considerado.

* Función de probabilidad

Si $\lambda = n r$

$r = \frac{\lambda}{n}$

$$p(x) = \lim_{n\to\infty} \binom{n}{x} r^x (1-r)^{n-x} = \lim_{n\to\infty} \binom{n}{x} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} =$$

$$= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} =$$

$$= \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \frac{\lambda^x}{n^x} \left(1 + \frac{1}{\left(-\frac{n}{\lambda}\right)}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} =$$

$$= \lim_{n\to\infty} \frac{n^x}{x!} \frac{\lambda^x}{n^x} \left[\left(1 + \frac{1}{\left(-\frac{n}{\lambda}\right)}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} = \boxed{\frac{\lambda^x}{x!} e^{-\lambda} \Rightarrow \text{distribución de Poisson}} \quad x \sim Pois[\lambda]$$

$$e^a = \sum_{n=0}^{\infty} \frac{a^n}{n!}$$

mar

* Función generadora de momentos

$$M_x(t) = \sum_{x=0}^{\infty} \frac{e^{xt} \lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

$$M'_x(t) = \lambda e^t e^{\lambda(e^t-1)}$$

$$M''_x(t) = \lambda e^t e^{\lambda(e^t-1)} + \lambda^2 e^{2t} e^{\lambda(e^t-1)} = \lambda e^t e^{\lambda(e^t-1)} (1 + \lambda e^t)$$

$$\mu = M'_x(0) = \lambda$$

$$\sigma^2 = M''_x(0) - M'_x(0)^2 = \lambda(1+\lambda) - \lambda^2 = \lambda$$

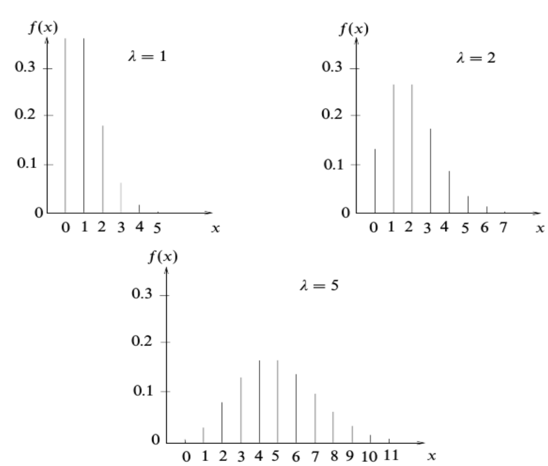


Figure 30.12 Three Poisson distributions for different values of the parameter λ.

Distribución hipergeométrica

- Experimento aleatorio: tenemos un conjunto de N bolas, M de ellas son rojas y $N-M$ blancas. Extraemos sucesivamente n bolas sin reposición y observamos el color de cada bola extraída, que puede ser rojo o blanco

- Espacio muestral $1 \leq n \leq N$, $0 \leq M \leq N$

$$\underbrace{BBB\cdots B}_{n} \qquad \underbrace{BB\cdots B}_{n-1} R \qquad \underbrace{BB\cdots B}_{n-2}\underbrace{RR}_{2} \qquad \underbrace{BB\cdots B}_{n-x}\underbrace{R\cdots R}_{x} \;\cdots$$

$$\uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow$$

$$C^{N-M}_{n} \qquad C^{N-M}_{n-1} \qquad C^{N-M}_{n-2} \times C^{M}_{2} \qquad C^{N-M}_{n-x} \times C^{M}_{x}$$

$$\downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow$$

- Variable aleatoria: nº de bolas rojas extraídas $x: \left(\operatorname{Max}\{0, n-(N-M)\}, 1, 2, 3, \ldots, \operatorname{Min}\{n, M\}\right)$

- Función de probabilidad = probabilidad de extraer x bolas rojas

Si $n \leq M$ entonces x)

$$p(x) = \frac{\text{nº de posibles formas de extraer } x \text{ bolas rojas}}{\text{nº de formas posibles de extraer } n \text{ bolas}} = \frac{C^{N-M}_{n-x} \times C^{M}_{x}}{C^{N}_{n}} =$$

$$= \frac{\binom{N-M}{n-x}\binom{M}{x}}{\binom{N}{n}}$$

Definimos $p = \frac{M}{N}$ $\quad q = 1-p \quad M = pN$

* Identidad de VanderMonde

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k} \Rightarrow \text{Generalización}$$

$$\binom{n_1+n_2+\dots n_p}{m} = \sum_{k_1+k_2+\dots k_p = m} \binom{n_1}{k_1}\binom{n_2}{k_2}\cdots\binom{n_p}{k_p}$$

* Valor medio: Si $0 = Max\{0, n-(N-M)\}$, $n = Min\{n, M\}$, entonces:

$$\mu = \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} x \binom{N-M}{n-x}\binom{M}{x}$$

Usar la identidad

$$x\binom{M}{x} = \frac{x\, M!}{x!\,(M-x)!} = \frac{M!}{(x-1)!\,(M-x)!} = \frac{M(M-1)!}{(x-1)!\,(M-x)!} = M\binom{M-1}{x-1}$$

$$\mu = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} M\binom{N-M}{n-x}\binom{M-1}{x-1} = \left[k = x-1\right] = \frac{M}{\binom{N}{n}} \sum_{k=0}^{n-1} \binom{N-M}{n-k-1}\binom{M-1}{k} \overset{\text{Usar}}{=} \text{Vandermonde}$$

$$= \frac{M}{\binom{N}{n}} \binom{N-1}{n-1} = \frac{M}{\frac{N!}{n!\,(N-n)!}} \frac{(N-1)!}{(n-1)!\,(N-n)!} = M \frac{n!\,\cancel{(N-n)}!}{N!} \frac{(N-1)!}{(n-1)!\,\cancel{(N-n)}!} =$$

$$= M\frac{n}{N} = np$$

* Varianza y desviación típica: Si $0 = Max\{0, n-(N-M)\}$, $n = Min\{n, M\}$, entonces:

$$E(x^2) = \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} x^2 \binom{N-M}{n-x}\binom{M}{x}.$$ sabemos que $x\binom{M}{x} = M\binom{M-1}{x-1}$

$$x(x-1)\binom{M}{x} = M(x-1)\binom{M-1}{x-1} = M\binom{M-1}{x-2} \Rightarrow x^2\binom{M}{x} - x\binom{M}{x} = M\binom{M-1}{x-2}$$

$$\Rightarrow x^2\binom{M}{x} = \binom{M-1}{x-1} + M\binom{M-1}{x-2} \Rightarrow$$

$$E(x^2) = \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1}\binom{M}{1} + \sum_{x=2}^{n} x^2 \binom{N-M}{n-x}\binom{M}{x} \right] =$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1}\binom{M}{1} + \sum_{x=2}^{n} \left[\binom{M-1}{x-1} + M\binom{M-1}{x-2} \right] \binom{N-M}{n-x} \right] =$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1} M + \sum_{x=2}^{n} \binom{M-1}{x-1}\binom{N-M}{n-x} + M \sum_{x=2}^{n} \binom{M-1}{x-2}\binom{N-M}{n-x} \right] =$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1} M - \binom{N-M}{n-1} + \sum_{x=1}^{n} \binom{M-1}{x-1}\binom{N-M}{n-x} + M \sum_{x=2}^{n} \binom{M-1}{x-2}\binom{N-M}{n-x} \right] = \begin{bmatrix} x-1=k \\ x-2=k \end{bmatrix}$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1}(M-1) + \sum_{k=0}^{n-1} \binom{M-1}{k}\binom{N-M}{n-k} + M \sum_{k=0}^{n-2} \binom{M-1}{k}\binom{N-M}{n-k} \right] =$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1}(M-1) + \sum_{k=0}^{n} \binom{M-1}{k}\binom{N-M}{n-k} + M \sum_{k=0}^{n} \binom{M-1}{k}\binom{N-M}{n-k} - \binom{M-1}{n}\binom{N-M}{0} - M\binom{M-1}{n-1}\binom{N-M}{1} - M\binom{M-1}{n}\binom{N-M}{1} \right]$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1}(M-1) - \cancel{\binom{N-1}{n}} - M(N-M)\left(\binom{M-1}{n-1} - \binom{M-1}{n} \right) + \cancel{\binom{N-1}{n}} + M\binom{N-1}{n} \right] =$$

$$= \frac{1}{\binom{N}{n}} \left[\binom{N-M}{n-1}(M-1) - M(N-M)\left(\binom{M-1}{n-1} - \binom{M-1}{n} \right) + M\binom{N-1}{n} \right] = ??$$

Resultado final: $, \sigma^2 = \frac{N-n}{N-1} n \frac{M}{N} \frac{N-M}{N}.$

Distribuciones continuas de probabilidad

$N(\mu,\sigma) =$

Distribution	Probability law $f(x)$	MGF	$E[X]$	$V[X]$
Gaussian	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	$\exp(\mu t + \frac{1}{2}\sigma^2 t^2)$	μ	σ^2
exponential	$\lambda e^{-\lambda x}$	$\left(\frac{\lambda}{\lambda - t}\right)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
gamma	$\frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}$	$\left(\frac{\lambda}{\lambda - t}\right)^r$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$
chi-squared	$\frac{1}{2^{n/2}\Gamma(n/2)}x^{(n/2)-1}e^{-x/2}$	$\left(\frac{1}{1-2t}\right)^{n/2}$	n	$2n$??
uniform	$\frac{1}{b-a}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$

Table 30.2 Some important continuous probability distributions.

* **Distribución gaussiana**: La distribución gaussiana es el límite de la binomial cuando el número total n de experimentos aleatorios elementales tiende a infinito manteniendo la probabilidad de éxito r finita.

$$p(x) = \binom{n}{x} r^x (1-r)^{n-x} \qquad \lim_{\substack{n\to\infty \\ x\to\infty}} \binom{n}{x} r^x (1-r)^{n-x} = \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{n!}{x!(n-x)!} r^x (1-r)^{n-x}$$

Utilizar la aproximación de Stirling $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, $n \gg 1$

$$\Rightarrow p(x) = \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n (e)^{n-x} r^x (1-r)^{n-x}}{\sqrt{2\pi x}\left(\frac{x}{e}\right)^x \sqrt{2\pi(n-x)}\,(n-x)^{n-x}} =$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}}\left(\frac{n}{x}\right)^{\frac{1}{2}} \frac{n^n}{x^x} \cancel{e^{x-n}} \cdot \cancel{e^{n-x}} (n-x)^{-\frac{1}{2}-n+x} r^x (1-r)^{n-x} =$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}} \frac{n^{n+\frac{1}{2}}}{x^{x+\frac{1}{2}}} (n-x)^{x-n-\frac{1}{2}} r^x (1-r)^{n-x} = \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}} \frac{n^{n+\frac{1}{2}}}{x^{x+\frac{1}{2}}} \frac{n^x}{n^x} (n-x)^{x-n-\frac{1}{2}} r^x (1-r)^{n-x}$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}} \frac{n^{x+\frac{1}{2}}}{x^{x+\frac{1}{2}}} \frac{n^n}{n^x} (n-x)^{x-n-\frac{1}{2}} r^x (1-r)^{n-x} =$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}} \left(\frac{x}{n}\right)^{-x-\frac{1}{2}} n^{n-x} (n-x)^{x-n-\frac{1}{2}} r^x (1-r)^{n-x} =$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}} \left(\frac{x}{n}\right)^{-x-\frac{1}{2}} n^{n-x} (n-x)^{x-n-\frac{1}{2}} r^x (1-r)^{n-x} \frac{n^{\frac{1}{2}}}{n^{\frac{1}{2}}} =$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi}} \left(\frac{x}{n}\right)^{-x-\frac{1}{2}} \frac{(n-x)^{x-n-\frac{1}{2}}}{n^{x-n-\frac{1}{2}}} \frac{r^x (1-r)^{n-x}}{n^{\frac{1}{2}}} =$$

$$= \lim_{\substack{n\to\infty \\ x\to\infty}} \frac{1}{\sqrt{2\pi n}} \left(\frac{x}{n}\right)^{-x-\frac{1}{2}} \left(1-\frac{x}{n}\right)^{x-n-\frac{1}{2}} r^x (1-r)^{n-x} =$$

$$= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi n}} \exp\left[-\left(x+\frac{1}{2}\right)\log\left(\frac{x}{n}\right) + \left(x-n-\frac{1}{2}\right)\log\left(1-\frac{x}{n}\right) + x\log r + (n-x)\log(1-r)\right]$$

hacemos el cambio $y = x - nr \Rightarrow x = y + nr$

$$= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi n}} \exp\left[-\left(y+nr+\frac{1}{2}\right)\log\left(r+\frac{y}{n}\right) + \left(y+n(r-1)-\frac{1}{2}\right)\log\left(1-r-\frac{y}{n}\right) + (y+nr)\log r + \right.$$

$$\left. + (n(1-r)-y))\log(1-r)\right] =$$

$$= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi n}} \exp\left[-\left(y+nr+\frac{1}{2}\right)\log\left[r\left(1+\frac{y}{nr}\right)\right] + \left(y-n(1-r)-\frac{1}{2}\right)\log\left[(1-r)\left(1-\frac{y}{n(1-r)}\right)\right] + \right.$$

$$\left. + (y+nr)\log r + (n(1-r)-y)\log(1-r)\right]$$

$$= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi n}} \exp\left[-\frac{1}{2}\log r - \frac{1}{2}\log(1-r) - \left(y+nr+\frac{1}{2}\right)\log\left(1+\frac{y}{nr}\right) + \left(y-n(1-r)-\frac{1}{2}\right)\log\left[1-\frac{y}{n(1-r)}\right]\right]$$

Hacer un desarrollo en serie del logaritmo y quedarse con los primeros términos

$$= \lim_{n\to\infty} \underbrace{\frac{1}{\sqrt{2\pi n r(1-r)}} \exp\left[-\frac{(x-nr)^2}{2nr(1-r)}\right]}_{\text{perfil gaussiano}} \Rightarrow \boxed{\begin{array}{c} B(n,r) \longrightarrow N\left(nr, \sqrt{nr(1-r)}\right) \\ n\to\infty \end{array}}$$

con $\mu = nr$ y $\sigma = \sqrt{nr(1-r)} \Rightarrow$ que son la media y la desviación típica de la binomial.

Aproximación gaussiana de la distribución de Poisson.

Distribución de Poisson $\Rightarrow p(x) = e^{-\lambda}\dfrac{\lambda^x}{x!}$

Vamos a ver que si $\lambda \to \infty$ entonces es posible aproximar la distribución de Poisson por una distribución normal:

$$p(x) = e^{-\lambda}\frac{\lambda^x}{x!} \Rightarrow \ln p(x) = -\lambda + x\ln\lambda - \ln x!$$

Aproximación de Stirling $\Rightarrow \begin{matrix}\lambda \gg 1 \\ \Downarrow \\ x \gg 1\end{matrix} \Rightarrow x! \sim (2\pi x)^{\frac{1}{2}} x^x e^{-x} \Rightarrow$

$$\Rightarrow \ln x! \sim \frac{1}{2}\ln(2\pi x) + x\ln x - x \Rightarrow$$

$$\Rightarrow \ln p(x) = -\lambda + x\ln\lambda - \frac{1}{2}\ln(2\pi x) - x\ln x + x$$

Hacer el cambio $\epsilon = x - \lambda$

$x = \epsilon + \lambda \Rightarrow x \gg 1 \Rightarrow \lambda \gg 1$ y $\epsilon \sim \lambda^{\frac{1}{2}}$

$$\Rightarrow \ln p(x) = -\lambda + (\epsilon+\lambda)\ln\lambda - \frac{1}{2}\ln(2\pi(\epsilon+\lambda)) - (\epsilon+\lambda)\ln(\epsilon+\lambda) + \epsilon + \lambda$$

$$= -\lambda + (\lambda+\epsilon)\left[\ln\lambda - \ln\left(\lambda\left(1+\frac{\epsilon}{\lambda}\right)\right)\right] + \lambda + \epsilon - \ln(2\pi(\epsilon+\lambda))^{\frac{1}{2}}$$

$$= -\lambda + (\lambda+\epsilon)\left[-\ln\left(1+\frac{\epsilon}{\lambda}\right) + 1\right] - \ln\left(2\pi(\lambda+\epsilon)\right)^{\frac{1}{2}}$$

// Desarrollamos en serie y nos quedamos con los primeros términos

$$= -\lambda + (\lambda+\epsilon)\left[1 - \ln\left(1+\frac{\epsilon}{\lambda}\right)\right] - \ln\left[2\pi\lambda\left(1+\frac{\epsilon}{\lambda}\right)\right]^{\frac{1}{2}} \Rightarrow$$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} \dots$$

$$\sim -\lambda + (\lambda+\epsilon)\left[-\frac{\epsilon}{\lambda} + \frac{\epsilon^2}{2\lambda^2} + 1\right] - \frac{1}{2}\ln(2\pi\lambda) - \frac{1}{2}\left(\frac{\epsilon}{\lambda} - \frac{\epsilon^2}{2\lambda^2}\right)$$

$$\sim -\cancel{\lambda} - \cancel{\epsilon} + \frac{\epsilon^2}{2\lambda} + \cancel{\lambda} - \frac{\epsilon^2}{\lambda} + \cancel{\epsilon} - \frac{1}{2}\ln(2\pi\lambda) - \frac{\epsilon}{2\lambda} + \frac{\epsilon^2}{4\lambda^2}$$

$$= -\frac{\epsilon}{2\lambda} + \frac{\epsilon^2}{4\lambda^2} - \frac{\epsilon^2}{2\lambda} - \frac{1}{2}\ln(2\pi\lambda).$$

Usamos que $\epsilon \sim \lambda^{\frac{1}{2}}$ para despreciar los 2 primeros términos

$$\Rightarrow \ln p(x) \sim -\frac{\epsilon^2}{2\lambda} - \frac{1}{2}\ln(2\pi\lambda)$$

$$p(x) = \frac{1}{\sqrt{2\pi\lambda}}\exp\left[-\frac{1}{2\lambda}(x-\lambda)^2\right] \Rightarrow$$

$$\boxed{\begin{matrix}\mathrm{Pois}[\lambda] \to N[\lambda, \lambda^{\frac{1}{2}}] \\ \lambda \to \infty\end{matrix}}$$

* Algunas propiedades relevantes de la distribución normal

$N(\mu,\sigma) \Rightarrow p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$ $\quad X \sim N(\mu,\sigma) \Rightarrow$ una variable aleatoria sigue una distribución de probabilidad normal.

$x = \mu \pm \sigma$ son puntos de inflexión de la curva que representa la distribución

* Distribución gaussiana estándar $N(0,1)$ $\quad \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$ $\quad z$ = variable aleatoria estándar

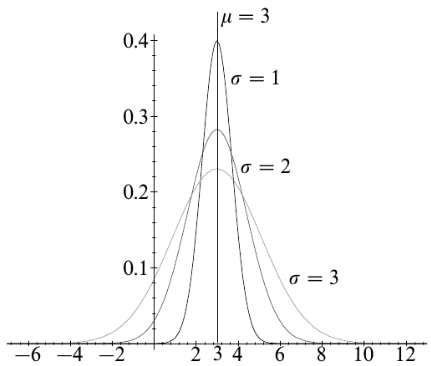


Figure 30.13 The Gaussian or normal distribution for mean $\mu = 3$ and various values of the standard deviation σ.

* Función de probabilidad acumulada

$$\Pr(X<x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} du \, \exp\left[-\frac{(u-\mu)^2}{2\sigma^2}\right] = \left[\begin{matrix} \frac{u-\mu}{\sigma} = z \\ du = \sigma\, dz \end{matrix}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} dz \, \exp\left[-\frac{z^2}{2}\right] = \Phi\left[\frac{x-\mu}{\sigma}\right]$$

donde $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} du \, e^{-\frac{u^2}{2}}$. Así que $\Pr(a \leq X \leq b) = \Phi\left[\frac{b-\mu}{\sigma}\right] - \Phi\left[\frac{a-\mu}{\sigma}\right]$

$\Phi(z) = \frac{1}{2} \mathrm{Erfc}\left[-\frac{x}{\sqrt{2}}\right]$ // $\mathrm{Erfc}(z) = 1 - \mathrm{Erf}(z)$, $\mathrm{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$

►*If X is described by a Gaussian distribution of mean μ and variance σ^2, calculate the probabilities that X lies within 1σ, 2σ and 3σ of the mean.*

$$\Phi(-n) = 1 - \Phi(n)$$

$$P_r(\mu - n\sigma \leq X \leq \mu + n\sigma) = \Phi\left(\frac{\mu + n\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - n\sigma - \mu}{\sigma}\right) = \Phi(n) - \Phi(-n) = 2\Phi(n) - 1$$

$n = 1 \Rightarrow 2\Phi(1) - 1 \approx 0.682689 \approx 68.3\%$

$n = 2 \Rightarrow 2\Phi(2) - 1 \approx 0.9545 \approx 95.4\%$

$n = 3 \Rightarrow 2\Phi(3) - 1 \approx 0.9973 \approx 99.7\%$

►*Sawmill A produces boards whose lengths are Gaussian distributed with mean 209.4 cm and standard deviation 5.0 cm. A board is accepted if it is longer than 200 cm but is rejected otherwise. Show that 3% of boards are rejected.*

Sawmill B produces boards of the same standard deviation but of mean length 210.1 cm. Find the proportion of boards rejected if they are drawn at random from the outputs of A and B in the ratio 3 : 1.

Sawmill A $\quad X \sim N(209.4, 5.) \quad P_r(X < 200\text{ cm}) = 0.030054 \approx 3.0\%$

Sawmill B $\quad X \sim N(210.1, 5.) \quad P_r(X < 200\text{ cm}) = 0.0216917 \approx 2.2\%$

* hay 4 líneas de distribución, 3 de A y 1 de B. Elegimos una línea al azar y después sacamos una plancha. ¿Cuál es la probabilidad de que la plancha sea rechazada?

$\frac{1}{4}$ Saco de A $\rightarrow p(\text{plancha rechazada}) = 0.03$

$\frac{1}{4}$ Saco de A $\rightarrow p(\text{plancha rechazada}) = 0.03$

$\frac{1}{4}$ Saco de A $\rightarrow p(\text{plancha rechazada}) = 0.03$

$\frac{1}{4}$ Saco de B $\rightarrow p(\text{plancha rechazada}) = 0.022$

$$p = \frac{3 \cdot 0.03 + 0.022}{4} = 0.028$$

Teorema del límite central

* Si tenemos un conjunto de variables aleatorias $x_1, \dots, x_n$ entonces tiene sentido definir su "combinación lineal":

$A_1 \equiv$ Suceso del espacio muestral E_1 que define la variable aleatoria x_1

$\vdots$

$A_n \equiv$ suceso del espacio muestral E_n que define la variable aleatoria x_n

$E = E_1 \times \dots \times E_n \Rightarrow$ espacio muestral que definiría la variable aleatoria $Z = c_1 X_1 + \dots + c_n X_n$

Si $A \in E$ entonces $A = (A_1, A_2, \dots, A_n)$

Definimos:

$$Z(A) = (c_1 x_1 + c_2 x_2 + \dots + c_n x_n)(A) = c_1 x_1(A_1) + c_2 x_2(A_2) + \dots + c_n x_n(A_n)$$

$c_1, \dots c_n \in \mathbb{R}$

* Si las variables aleatorias $\{x_i\}$ son continuas entonces Z también será una variable aleatoria continua.

<u>Teorema del límite central</u>: Sea $X_i, i = 1, \dots n$ un conjunto de variables aleatorias independientes, cada una de ellas con su función densidad de probabilidad $f_i(x)$ (no necesariamente iguales) con media μ_i y varianza σ_i^2. La variable aleatoria

$Z \equiv \frac{1}{n} \sum_{i=1}^{n} X_i$ tiene las siguientes propiedades:

i) El valor medio de su distribución de probabilidad es $E(Z) = \frac{1}{n} \sum_{i=1}^{n} \mu_i$

ii) La varianza de su distribución de probabilidad es $V(Z) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2$

iii) Finalmente se tiene que si $p_n(z)$ es la distribución de probabilidad asociada a Z y los recorridos de las variables X_i y Z coinciden con la recta real entonces $\lim_{n\to\infty} p_n(z) = N(\mu, \sigma)$, $\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i$, $\sigma^2 = \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2$

Prueba: Si las variables aleatorias $\{x_i\}$ se corresponden con experimentos aleatorios independientes entonces la función densidad de probabilidad conjunta viene dada por

$$f(x_1,\dots x_n) = f_1(x_1)\cdots f_n(x_n) \quad \cdot I = Z(I_1\times\dots\times I_n) \quad \begin{array}{l}(I \text{ recorrido de } Z)\\ (I_k \text{ recorrido de } x_k)\end{array}$$

i) $\mu = E(Z) = \int_I z\, p_n(z)\, dz = \int_{I_1\times\dots\times I_n} \frac{(x_1+\dots+x_n)}{n} f(x_1,\dots x_n)\, dx_1\dots dx_n =$

$$\int_{I_1\times\dots\times I_n} \frac{(x_1+\dots+x_n)}{n} f_1(x_1)\cdots f_n(x_n)\, dx_1\dots dx_n =$$

$$= \frac{1}{n}\left[\underbrace{\left(\int_{I_1} x_1 f(x^1)\, dx^1\right)}_{\mu_1} + \dots + \underbrace{\left(\int_{I_n} x_n f(x^n)\, dx^n\right)}_{\mu_n}\right] = \frac{1}{n}\left[\mu_1 + \dots + \mu_n\right] = \frac{1}{n}\sum_{i=1}^n \mu_i$$

ii) $\sigma^2 = E(Z^2) - E(Z)^2 = E(Z^2) - \mu^2 \quad \| \quad \sigma_i^2 = E(x_i^2) - E(x_i)^2 = E(x_i^2) - \mu_i^2$

$$E(Z^2) = \int_I z^2 p_n(z)\, dz \quad = \int_{I_1\times\dots\times I_n} \frac{(x_1+\dots+x_n)^2}{n^2} f(x_1,\dots,x_n)\, dx_1\dots dx_n =$$

$$= \frac{1}{n^2}\int_{I_1\times\dots\times I_n} \sum_{i,j}^n x_i x_j f_1(x_1)\cdots f_n(x_n)\, dx_1\dots dx_n = \frac{1}{n^2}\sum_{\substack{i,j\\ i\neq j}}^n \int_{I_i} x_i f_i(x_i)\, dx_i \int_{I_j} x_j f_j(x_j)\, dx_j +$$

$$+ \frac{1}{n^2}\sum_{i=1}^n \int_{I_i} x_i^2 f(x_i)\, dx_i = \frac{1}{n^2}\sum_{\substack{i,j\\ i\neq j}}^n \mu_i\mu_j + \frac{1}{n^2}\sum_{i=1}^n E(x_i^2) =$$

$$= \frac{1}{n^2}\left(\sum_{\substack{i,j\\ i\neq j}}^n \mu_i\mu_j + \sum_{i=1}^n \sigma_i^2 + \mu_i^2\right) = \frac{1}{n^2}\left(\sum_{i,j}^n \mu_i\mu_j + \sum_{i=1}^n \sigma_i^2\right) \Rightarrow \text{sustituyendo}$$

$$\sigma^2 = \frac{1}{n^2}\left(\sum_{i,j}^n \mu_i\mu_j + \sum_{i=1}^n \sigma_i^2\right) - \frac{1}{n^2}\left(\sum_{i=1}^n \mu_i\right)^2 = \sum_{i=1}^n \sigma_i^2$$

iii) Prueba abreviada: Calcular $M_Z(t)$

$$M_Z(t) = \int_I e^{zt} p_n(z)\, dz = \int_{I_1\times\cdots\times I_n} e^{(x_1+\cdots+x_n)\frac{t}{n}} f_1(x_1)\cdots f_n(x_n)\, dx_1\cdots dx_n =$$

$$= \left[\int_{I_1} e^{\frac{t x_1}{n}} f_1(x_1)\, dx_1\right] \cdots \left[\int_{I_n} e^{\frac{t x_n}{n}} f_n(x_n)\, dx_n\right] = \prod_{i=1}^{n} M_{X_i}\left(\frac{t}{n}\right)$$

Ahora bien, si $n \gg 1$

$$M_{X_i}\left(\frac{t}{n}\right) = \int_{I_n} e^{t\frac{x_i}{n}} f_i(x_i)\, dx_i \simeq \int_{I_n} \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{t\, x_i}{n}\right)^k f_i(x_i)\, dx_i =$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \frac{t^k}{n^k} E(x_i^k)$$

$$\simeq 1 + \frac{t}{n}\mu_i + \frac{t^2}{2n^2} E(x_i^2) = 1 + \frac{t\mu_i}{n} + \frac{t^2}{2n^2}(\sigma_i^2 + \mu_i^2)$$

ge

Esto coincide con el desarrollo asintótico de $M_{X_i}\left(\frac{t}{n}\right) \approx \exp\left(\frac{\mu_i t}{n} + \frac{1}{2}\sigma_i^2 \frac{t^2}{n^2}\right)$, cuando $n \to \infty$

Lo comprobamos para los primeros términos $\Rightarrow$ $e^{\frac{\mu_i t}{n}} e^{\frac{1}{2}\sigma_i^2 \frac{t^2}{n^2}} =$

$$= \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{\mu_i t}{n}\right)^k \sum_{q=0}^{\infty} \frac{1}{2^q} \sigma_i^{2q} \frac{t^{2q}}{n^{2q}} \frac{1}{q!} = \left(1 + \frac{\mu_i t}{n} + \frac{\mu_i^2 t^2}{2n^2} + \cdots\right)\left(1 + \frac{\sigma_i^2 t^2}{2n^2} + \cdots\right.$$

$$\simeq \left(1 + \frac{t\mu_i}{n} + \frac{t^2}{2n^2}(\mu_i^2 + \sigma_i^2)\right) + \cdots$$

Por tanto $\lim_{n\to\infty} M_Z(t) = \lim_{n\to\infty} \prod_{i=1}^{n} \exp\left[\frac{\mu_i t}{n} + \frac{1}{2}\sigma_i^2 \frac{t^2}{n^2}\right] =$

$$= \lim_{n\to\infty} \exp\left[t \sum_{i=1}^{n} \frac{\mu_i}{n} + \frac{t^2}{n^2} \sum_{i=1}^{n} \sigma_i^2 \right] = \exp\left[t\mu + t^2\sigma^2 \right]$$

$$\left.\begin{aligned} \sigma^2 &= \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 \\ \mu &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mu_i \end{aligned}\right\}$$ Así que $p(z) = \lim_{n\to\infty} p_n(z)$ cumple que

$$\int_I e^{tz} p(z)\,dz = \exp\left[t\mu + t^2\sigma^2 \right]$$

Si $I = \infty \Rightarrow$ $p(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\Rightarrow$ distribución gaussiana. Q.E.D.

* Ejemplo: Lanzamos una piedra y estudiamos donde cae. Aunque teóricamente podemos calcular de manera exacta el punto donde cae la piedra, en la práctica la trayectoria de la piedra está afectada por variables que desconocemos (velocidad del viento, rozamiento de la piedra con el aire, incertidumbre en la velocidad inicial, ... Así que podemos considerar que "tirar una piedra y ver donde cae" es un experimento aleatorio. Variable aleatoria $\Rightarrow$ posición del punto de impacto de la piedra. Suponemos que es unidimensional y equivalente a la recta real

$\mathbb{R}$

En general la distribución de probabilidad $f_i(x_i)$ asociada a la tirada i es desconocida pero si definimos una nueva variable aleatoria $Z = \frac{1}{n} \sum_{i=1}^{n} x_i$ entonces la función densidad de probabilidad $p_n(Z)$ que resulta de hacer un cambio de variables aleatorias en $f_1(x_1) \dots f_n(x_n)$ se podrá aproximar por una normal si $n \gg 1$