Estadística

Todo lo necesario para no cagarla

Universitat d'Alacant
Universidad de Alicante

Resumen

Veremos la teoría, como aplicarla y ejemplos para el contenido de Estadística .

Eduardo Espuch

Curso 2019-2020

Índice

1. Combinatoria

El análisis combinatorio permite que, a partir de un conjunto con m elementos, calcular el numero de subconjuntos que cumplen unos requisitos preestablecidos. Considerando m como la cantidad de elementos en total de un conjunto y n la cantidad de elementos en el subconjuntos que se quiere estudiar, tenemos los siguientes requisitos:

- ¿Es n=m?

  En el caso de que se cumpla, tratamos con una **Permutación**

  En el caso de que no se cumpla(porque m>n, tenemos que considerar:

- ¿Importa el orden?

  En el caso afirmativo, tratamos con **Variaciones**

  En el caso negativo, tratamos con **Combinaciones**

1.1. Permutaciones de orden m

Llamaremos permutación a la reorganización de un conjunto de elementos. Al obtener la cantidad de permutaciones posibles en un conjunto de m elementos se obtendrá el numero de distintas formas en las que estos se pueden ordenar, el propio conjunto se consideraría como una permutacion.

En el caso de haber repeticiones, es decir, que de los m elementos, al menos existe un par de ellos que son idénticos, se deberán de considerar representantes de clase para los elementos, asignando en estos la cantidad de elementos que se repiten por clase.

Las permutaciones de orden m sin repeticiones se denotan con P_m y con repetición se denotarían con $PR_m^{a_1,a_2,\ldots,a_r}$ siendo m el numero total de elementos del conjunto y $a_i, 1 \leq i \leq r$ los representantes, donde se calculan con:

Formulas Permutaciones

$$P_m = m! = 1{\cdot}2{\cdot}...{\cdot}m \qquad\qquad PR_m^{a_1,a_2,\ldots,a_r} = \frac{m!}{a_1!, a_2!, ..., a_r!}$$

1.2. Variación de orden n

Llamaremos variación al subconjunto ordenado de n elementos formado a partir de los m elementos del conjunto. Al querer obtener las variaciones de un conjunto, obtenemos todos los distintos subconjuntos ordenados que se pueden formar a partir de este.

Al tratar con variaciones sin repeticiones, esto no significa que no puedan haber al menos un par elementos idénticos dentro del conjunto, únicamente que un mismo elemento no podrá repetirse en el subconjunto, es decir, dados a y b dentro de un conjunto, y con valor idéntico, no se considerarían repetición. Si consideramos variaciones con repetición, tenemos que considerar m como el numero de representantes de clases, y, al desconocer el numero de elementos existentes para cada clase, los consideramos ilimitados. Es la forma apropiada para mostrar todas las posibles formas de ordenar en un subconjunto de n elementos

Las variaciones de orden n sin repeticiones se denotan con V_m^n o $V_{m,n}$ y con repetición se denotarían con VR_m^n o $VR_{m,n}$ siendo m el numero total de elementos del conjunto y n, donde se calculan con:

Formulas Variaciones

$$V_m^n = V_{m,n} = \frac{m!}{(m-n)!} \qquad VR_m^n = VR_{m,n} = m^n$$

1.3. Combinación de orden n

Llamaremos combinación al subconjunto no ordenado de n elementos formado a partir de los m elementos del conjunto. Al querer obtener las combinaciones de un conjunto, obtenemos todos los distintos subconjuntos que se pueden formar a partir de este.

Al tratar con repeticiones, se considera igual que con las variaciones, donde m es para los representantes y existen un ilimitado numero de elementos para dichos representantes.

Las combinaciones de orden n sin repeticiones se denotan con C_m^n o $c_{m,n}$ y con repetición se denotarían con cR_m^n o $cR_{m,n}$ siendo m el numero total de elementos del conjunto y n, donde se calculan con:

Formulas Combinaciones

$$C_m^n = C_{m,n} = \begin{pmatrix} m \\ n \end{pmatrix} = \frac{m!}{n!(m-n)!} \qquad CR_m^n = CR_{m,n} = \begin{pmatrix} m+n-1 \\ n \end{pmatrix} = \frac{(m+n-1)!}{n!(m-1)!}$$

Anotaciones mías

Puede darse el caso de un conjunto que requiere de varias condiciones, esto puede indicarse usando un superconjunto que engloba al resto, tal que cada uno hace referencia a una condicion. Denotando S como al superconjunto y siendo $C_i, \forall i \in I$ los conjuntos que cumplen distintas condiciones, se tendrá que $S = \{C_i | \forall i \in I\}$.

Es mas, estos conjuntos que cumplen distitnas condiciones, pueden tratarse con combinaciones, permutaciones o varaiciones, enlazandose usando cuna conjuncion (operador 'y'), lo que se traduce a usar una multiplicacion.

Por ejemplo, se pide obtener las distintas formas de sentar 5 personas si sabemos que solo 2 pueden conducir.

Considerando C_1 := Pilotos y C_2 :=Pasajeros, junto con S :=Ocupantes del vehículo, tendremos que obtener las posibles opciones para pilotos y para pasajeros para obtener las distintas formas de ordenar a los ocupantes del vehiculo considerando las condiciones dadas.

- Obtener los posibles pilotos. Sabiendo que solo hay dos pilotos pero por motivos mas que obvios, solo uno puede conducir. No pdoremos usar Permutaciones. Ademas, como solo tratamos con subconjuntos de orden 1, el orden dara igual, por lo tanto podremos trabajar con Combinaciones.

$$C_{2,1} = \frac{2!}{1!(2-1)!} = 2$$

- Obtener los posibles pasajeros. Sabiendo que se trataran con 4 ocupantes a la vez de los 4 restantes (considerando que uno de los 5 tiene que ser estrictamente el piloto), trabajaremos con Permutaciones.

$$P_4 = 4! = 24$$

- Los posibles ocupantes con la condición dada: una forma de verlo es que, para un piloto dado, hemos obtenido una cantidad de permutaciones posibles (24), y al tener 2 posibles pilotos, tendremos otro ser de permutaciones cambiando donde se sentaba el actual piloto con el anterior, teniendo asi $2 \cdot 24 = 48$ posibles.

2. Sucesos aleatorios

Al realziar un experimento sobre un fenomeno aleatorio, el conjunto de resultados impredecibles quedan guardados en el espacio muestral y llamaremos suceso a cada subconjunto que se pueda formar. Si los consideramos como conjuntos, tenemos las siguientes propiedades:

Ω : Espacio muestral $\emptyset$: Conjunto vacío S : Suceso

$card(S) = 1$ suceso elemental $card(S) > 1$ suceso compuesto

$card(S) = 0 \mapsto S = \emptyset$ suceso imposible $card(S) = card(\Omega)$ Suceso seguro

Si un experimento obtiene como resultado uno de los elementos de un suceso, estara verificado.

Las propiedades de los conjuntos pueden ser aplicados. A continuacion veremos conceptos relacionados con aplicar dichas propiedades.

Diremos que dos o mas conjuntos son **incompatibles** si la interseccion entre cada par de conjuntos es vacia,

$$A_1, A_2, ..., A_n \mapsto A_i \cap A_j = \emptyset | \forall i, j \in I, i \neq j$$

2.1. Funciones de probabilidad

Introducimos ahora las funciones de probabilidad, que denotaremos con $P(S)$ mezclando el análisis combinatorio con los fenómenos aleatorios tal que:

$$\begin{array}{l} P : \Omega \mapsto \mathbb{R}\} \text{casos posibles} \\ \left.\begin{array}{r} A \mapsto \mathbb{R} \\ B \mapsto \mathbb{R} \end{array}\right\} \text{casos favorables} \quad P(\Omega) = 1 > P(\emptyset) = 0 \end{array}$$

$$P(A) = \frac{\text{casos favorables de A}}{\text{casos posibles}}$$

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

o, si descomponemos...

$$P(A \cup B) = P(\overline{A} \cap B) + P(A \cap \overline{B}) + P(A \cap B)$$

Al tratar con varios sucesos la operacion se complica tal que:

$$P(A_1 \cup ... \cup A_n) = \sum_{i}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + ... + (-1)^{n+1} P(A_1 \cap ... \cap A_n)$$

Si se da que podemos interpretar los conjuntos graficamente, se pueden descomponer y sumar directamente.

Si sabemos que un conjunto de sucesos son incompatibles, sabemos que la interseccion entre cada par de sucesos es vacia y por lo tanto, se puede afirmar que se cumple $P(\bigcup_{i=i}^{n} A_i) = \sum_{i}^{n} P(A_i)$

Diremos que la probabilidad del suceso A esta condicionado al suceso B si para obtener la probabilidad de A debemos de considerar que se ha verificado B, es decir, que ocurra B es necesario para que ocurra A. Se denota con $P(A/B)$ y se calcula tal que:

$$P(A/B) = \frac{P(A \cap B}{P(B)} \mapsto P(A \cap B) = P(A/B)P(B)$$

Es mas, diremos que un conjunto de sucesos son **independientes** si $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$ donde $P(A_i) \neq 0, \forall i \in I$

2.2. Probabilidad total y Teoream de Bayes

Si se da el caso de que tenemos un conjunto de sucesos incompatibles entre si (interseccion vacia entre cada par) y cuya union forma todo el espacio muestral, hablaremos de que estos forman una particion sobre el espacio muestral.

La **Probabilidad total** es una formula en la que podemos obtener la probabilidad de un suceso si se tiene:

- Partición sobre el espacio muestral. Se conocen las probabilidades de todos los sucesos en la partición.
- La probabilidad del suceso deseado condicionado a cada uno de los sucesos de la partición.

Se calcula con $P(B) = \sum_{i=1}^{n} P(A_i)P(B/A_i)$

El Teorema de Bayes es una formula en la que, aplicando la probabilidad total junto con la propiedad conmutativa y la definicion de un suceso condicionado a otro, se puede obtener la probabilidad de un suceso parte de la particion condicionado a un suceso que puede ser desconocido. Para ello, se tiene que:

- Partición sobre el espacio muestral. Se conocen las probabilidades de todos los sucesos en la partición.
- La probabilidad del suceso deseado condicionado a cada uno de los sucesos de la partición.

Con la definición de suceso condicionado a otro, tenemos $P(A_k/B) = \dfrac{P(A_k \cap B)}{P(B)}$, al aplicar la conmutativa sobre la intersección tendríamos $P(B \cap A_k)$ y, al volver a aplicar la definicion, entonces obtenemos $P(B \cap A_k) = P(B/A_k)P(A_k)$, dejando la primera funcion a

$$P(A_k/B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B/A_k)P(A_k)}{P(B)}$$

Si, ademas, aplicamos la probabilidad total apra obtener $P(B)$, entonces se quedaria la formula del Teorema de Bayes tal que:

$$P(A_k/B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B/A_k)P(A_k)}{\sum_{i=1}^{n} P(A_i)P(B/A_i)}$$

Anotaciones mías

3. Variables aleatorias

Llamaremos variables aleatoria unidimensional a toda aplicacion que asigna un valor a un suceso tal que:

$$
\begin{aligned}
X : \Omega &\mapsto \mathbb{R} \\
A &\mapsto x_1 \\
B &\mapsto x_2 \\
X = \{x_1, x_2, ...\}
\end{aligned}
$$

De esta forma, podemos definir funciones sobre sucesos, distinguiendo dos tipos de conjuntos sobre los que trabajar, el conjunto de variables discreta (valores que suelen ser enteros y existira una cantidad finita) o de variables continua (valores que suelen ser reales y entre dos valores siempre puede existir uno).

Vamos a enlistar a continuación las distintas funciones unidimensionales:

Funciones de distribucion

Pueden usarse sobre: discreta y continua

Se denota con

$$F(x_k) = P(X \leq x_k) = \sum_{i}^{k} P(X = x_i),\ \forall x_i \in X \wedge i \in I$$

que se interpreta como la probabilidad acumulada hasta el valor indicado (x_k).

Y cumple las siguientes propiedades:

1. La probabilidad acumulada estará contenida entre 0 y 1.

$$0 \leq F(x) \leq 1,\ \forall x$$

2. La probabilidad acumulada para el menor posibible sera 0 y para el mayor posible sera 1, debido a que las variables aleatorias estan ordenadas.

$$
\begin{aligned}
F(+\infty) = 1 &\quad (P(X \leq +\infty) = P(\Omega) = 1) \\
F(-\infty) = 0 &\quad (P(X \leq -\infty) = P(\emptyset) = 0)
\end{aligned}
$$

3. $F(x)$ es no decreciente, es decir, si $a < b$ se tiene que dar que $F(a) \leq F(b)$

4. Dado el intervalo $[a, b]$, del cual se quiere obtener la probabilidad acumulada, entonces $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$

5. En el caso de querer obtener la probabilidad acumulada en el sentido opuesto (valores mayores que un valor) se considera tomar el complementario restandose al maximo valor posible acumulado (1). La formula se veria:

$$P(X > a) = 1 - P(X \leq a) = 1 - F(a)$$

6. $F(x)$ es continua por la derecha

Funciones de cuantía (o probabilidad)

Pueden usarse sobre: discreta

Se denota con

$$f(x) = P(X = x),\ \forall x$$

que se interpreta como la probabilidad de que se verifique un suceso. Al tratarse con un conjunto discreto (finito o infinito numerables) de valores, se observaria que graficamente es una funcion de escalera (se dan saltos al alncanzar cada punto).

Y cumple las siguientes propiedades:

1. La probabilidad de un suceso estará contenida entre 0 y 1.
$$0 \leq f(x) \leq 1, \, \forall x$$

2. La suma de todas las probabilidades de los sucesos dara como total 1, tal que
$$\sum_{x_i} f(x_i) = \sum_{x_i} P(X = x_i) = 1$$

   De esta forma, relacionamos las funciones de distribucion con la de cuantía.
$$F(x_k) = P(X \leq x_k) = \sum_{i}^{k} P(X = x_i) = \sum_{i}^{k} f(x_i), \, \forall x_i \in X \wedge i \in I$$

3. Diremos que trataremos con una variables discreta uniforme si cada suceso es igual de probable que el resto.

Funciones de densidad (o probabilidad)

Pueden usarse sobre: continua

Se denota con la función de distribución definida, siendo es continua, derivable y con derivada continua, siendo esta derivada la funcion de densidad, tal que
$$f(x) = \frac{dF(x)}{dx}$$

Y cumple las siguientes propiedades:

1. Para cada posible suceso, tendra una probabilidad igual o mayor que 0, no se espera que supere el 1, es decir, $f(x) \geq 0, \, \forall x$

2. La integral entre el minimo y maximo valor posible de la funcion de densidad sera basicamente obtener la probabilidad de un suceso seguro, tal que
$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

3. Como las formulas de densidad y distribucion estan relacionadas, algunas propiedades se ven reutilizadas, como puede ser el hecho de obtener la probabilidad para un intervalo de valores, teniendo que
$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) = [F(x)]_a^b = \int_a^b f(x)dx$$

4. Diremos que tratamos con una variable continua uniforme sobre un intervalo $[a, b]$ si la funcion de densidad en dicho intervalo es contante y nula fuera de este.

Habrá que repasar derivadas e integrales, revisar el anexo.

Llamaremos variable aleatoria bidimensional a toda aplicacion que asigna dos valores a un par de sucesos. Visto de otra manear, es un par ordenado de variables aleatorias unidimensionales donde cada una hace referencia a un suceso diferente. Esta aplicacion es:

$$\begin{aligned} (X, Y) : \Omega \mapsto \mathbb{R}^{\nvDash} & \\ A \mapsto (x_1, y_1) & \\ B \mapsto (x_2, y_2) & \\ (X, Y) = \{(x_1, y_1), (x_2, y_2,)...\} & \end{aligned}$$

Vamos a enlistar a continuación las distintas funciones bidimensionales

Funciones de distribucion

Pueden usarse sobre: discreta y continua

Se denota con

$$F(x_n, y_m) = P(X \leq x_n, Y \leq y_m) = \sum_i^n \sum_j^m P(X = x_i, Y = y_j),\ \forall (x_i, y_j) \in (X, Y) \wedge i, j \in I$$

que se interpreta como la probabilidad acumulada hasta el par de valores indicado (x_n, j_m). Se vera mas adelante que se definira con otras formulas en lugar del sumatorio de probabilidades directamente.

Y cumple las siguientes propiedades:

1. La probabilidad acumulada estará contenida entre 0 y 1.

$$0 \leq F(x, y) \leq 1,\ \forall (x, y)$$

2. La probabilidad acumulada para un par de valores con el mayor valor posible sera de 1, ya que agrupa el total de valores pero, la probabilidad acumulada para un par de valores donde uno es el menor valor posible, sera de 0, ya que no existirá ningún par de sucesos posible con esta combinacion.

$$\begin{array}{ll} F(+\infty, +\infty) = 1 & \quad (P(X \leq +\infty, Y \leq +\infty) = P(\Omega) = 1) \\ F(-\infty, y) = F(x, -\infty) = 0 & \end{array}$$

3. Si se desea obtener la probabilidad acumulada en un intervalo $[(a_1, b_1), (a_2, b_2)]$, similar que con la unidimensional, tendremos

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) =$$
$$P(X \leq a_2, Y \leq b_2) - P(X \leq a_2, Y \leq b_1) - P(X \leq a_1, Y \leq b_2) + P(X \leq a_1, Y \leq b_1) =$$
$$F(a_2, b_2) - F(a_2, b_1) - F(a_1, b_2) + F(a_1, b_1)$$

4. $F(x, y)$ es continua por la derecha

Funciones de cuantía (o probabilidad) conjunta

Pueden usarse sobre: discreta

Se denota con

$$f(x, y) = P(X = x, Y = y),\ \forall (x, y)$$

que se interpreta como la probabilidad de que se verifique un par de sucesos..

Y cumple las siguientes propiedades:

1. La probabilidad de un suceso estará contenida entre 0 y 1.

$$0 \leq f(x, y) \leq 1,\ \forall (x, y)$$

2. La suma de todas las probabilidades de los sucesos dara como total 1, tal que

$$\sum_i \sum_j f(x_i, y_j) = \sum_i \sum_j P(X = x_i, Y = y_j) = 1$$

De esta forma, relacionamos las funciones de distribucion con la de cuantía.

$$F(x_n, y_m) = P(X \leq x_n, Y \leq y_m) = \sum_i^n \sum_j^m P(X = x_i, Y = y_j) =$$

$$\sum_i^n \sum_j^m f(x_i, y_j),\ \forall (x_i, y_j) \in (X, Y) \wedge i, j \in I$$

Funciones de densidad (o probabilidad)

Pueden usarse sobre: continua

Se denota con la función de distribución definida, siendo es continua, derivable y con derivada continua, siendo esta derivada la funcion de densidad, tal que

$$f(x) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Y cumple las siguientes propiedades:

1. Para cada posible par de sucesos, tendra una probabilidad igual o mayor que 0, no se espera que supere el 1, es decir, $f(x, y) \geq 0,\ \forall (x, y)$
2. La integral entre el minimo y maximo valor posible de la funcion de densidad sera basicamente obtener la probabilidad de un suceso seguro, tal que
$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$
3. Como las formulas de densidad y distribucion estan relacionadas, algunas propiedades se ven reutilizadas, como puede ser el hecho de obtener la probabilidad para un intervalo de valores, teniendo que
$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{b_1}^{b_2} \int_{a_1}^{a_2} f(x, y) dx dy$$
Y, lo que es mas importante, obtener una definicion de la funcion de distribucion para variables bidimensionales continuas tal que:
$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t) ds dt$$

3.1. Distribuciones marginales

Usadas en variables bidimensionales, tanto discretas como continuas, las distitnas distribuciones marginales constan en tratar al entorno en funcion de una de los variables, es decir, se ignora un suceso para fijarse en uno exclusivamente. Veamos las tres funciones anteriores. Matematicamente es la forma de expresar una aplicacion que, tomando una espacio compuesto por dos sucesos, devuelve como resultado un valor de uno de los dos sucesos.

Funciones de distribucion marginal

Usado para discretas y continuas, distinguiremos a la funcion de distribucion marginal de X (Y) a la distribucion unidimensional que tiene X(Y), prescindiendo de Y(X).

$$F_1(x) = P(X \leq x) = P(X \leq x, Y \leq +\infty) = F(x, +\infty)$$

$$F_2(y) = P(Y \leq y) = P(X \leq +\infty, Y \leq y) = F(+\infty, y)$$

F_1 es como denotaremos a la distribucion marginal de X y F_2 a la de Y, por lo general, los subindices denotaran a la variable por lo tanto esta aclaracion no se realizara en futuras explicaciones.

Cumple las mismas propiedades que haria una funcion para variables unidimensionales, por lo tanto, veremos que se cumple tambien definiciones realcionadas a esta.

Funciones de cuantia (o probabilidad) marginal

Usado para discretas, se tiene que

$$f_1(x) = \sum_j f(x, y_j) = \sum_j P(X = x, Y = y_j), \quad \forall x$$

$$f_2(y) = \sum_i f(x_i, y) = \sum_i P(X = x_i, Y = y), \quad \forall y$$

Las mismas propiedades que para una variable unidimensional discreta, y consideramos que podemos obtener la distribucion marginal usando esta funcion de tal manera que(tomando la marginal de X como ejemplo):

$$F_1(x_n) = \sum_i^n \sum_j f(x_i, y_j) = \sum_i^n \sum_j P(X = x_i, Y = y_j) = P(X \leq x_n, Y \leq +\infty)$$

Funciones de densidad (o probabilidad) marginal

Usado para continuas, se tiene que

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)dy, \quad \forall x$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y)dx, \quad \forall y$$

Cumple las mismas propiedades que para una variable unidimensional y podemos definir a una funcion de distribucion marginal usando esta funcion de tal manera que (tomando la marginal de X como ejemplo):

$$F_1(x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{+\infty} f(x, y)dydx = P(X \leq x_n, Y \leq +\infty)$$

Esta funcion puede verse compleja, pero realmente consiste en obtener la que seria la funcion de densidad marginal de X y a continuacion obtener el valor que corresponderia de acumular probabilidades hasta el valor indicado (x_n).

En el caso de tener la funcion de distribucion marginal y querer obtener la funcion de densidad marginal, habria que hacer la derivada parcial y, si se desea obtener la funcion de densidad conjunta, entonces volver a derivar por la variable correspondiente.

Independencia de variables

Diremos que dos variables son independientes si se cumple que

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

es decir, si $F(x,y) = F_1(x)F_2(y),\ \ \forall(x,y) \in (X,Y)$, esto, volviendo al tema 2 donde definiamos las funciones de probabilidad sobre sucesos, teniamos que si la probabilidad de la interseccion de dos sucesos (dos sucesos se verifican a la vez) es igual al producto de la probabilidad de ambos sucesos por separado (que los sucesos se verifiquen sin considerar al otro), se afirma que son independientes.

Se recomienda que, para obtener si una variable bidimensional es independiente, comprobar si no es dependiente, es decir, si se encuentra un caso donde la igualdad anteriormente descrita no se cumple, entonces es dependiente y, de tal modo, se puede afirmar que no es independiente. La forma mas comoda de calcularlo es usando las funciones de probabilidad conjunta igualada al producto de las marginales.

3.2. Distribuciones condicionales

Vamos a observar ahora las funciones de probabilidad y distribucion para sucesos condicionados a otros. Vamos a destacar que estas funciones se dan con variables preestablecidades pero, en ocasiones (y generalemente sobre las de probabilidad) se espera obtener una tabla indicando el valor correspondiente de la variable condicionada a un suceso fijo).

Funcion de distribucion condicional

Dado x condicionado a y, tenemos:

$$F(x/y) = P(X \leq x/Y \leq y) = \frac{P(X \leq x, Y \leq y)}{P(Y \leq y)} = \frac{F(x,y)}{F_2(y)}$$

Dado y condicionado a x, tenemos:

$$F(y/x) = P(Y \leq y/X \leq x) = \frac{P(X \leq x, Y \leq y)}{P(X \leq x)} = \frac{F(x,y)}{F_1(x)}$$

Funcion de probabilidad condicional

Dado x condicionado a y, tenemos:

$$g_1(x/y) = P(X = x/Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x,y)}{f_2(y)},\ \text{opcional seria si } \forall x$$

x	x_1	$\dots$	x_n
$g_1(x_i/y)$		$\dots$	

Dado y condicionado a x, tenemos:

$$g_2(y/x) = P(Y = y/X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x,y)}{f_1(x)}, \text{ opcional seria si } \forall y$$

y	y_1	$\dots$	y_n
$g_2(y_i/x)$		$\dots$	

Funcion de densidad condicional

Dado x condicionado a y, tenemos:

$$g_1(x/y) = \frac{f(x,y)}{f_2(y)}, \text{ opcional seria si } \forall x$$

x	x_1	$\dots$	x_n
$g_1(x_i/y)$		$\dots$	

Dado y condicionado a x, tenemos:

$$g_2(y/x) = \frac{f(x,y)}{f_1(x)}, \text{ opcional seria si } \forall y$$

y	y_1	$\dots$	y_n
$g_2(y_i/x)$		$\dots$	

Anotaciones mias

Al calcular las funciones, en muchos casos puede ser continua en intervalos pero no continua en el conjunto de los reales. Obtener la funcion para distintos intervalos e indicarlo bien.

4. Esperanza y momentos

No trataremos momentos ni la desigualdad de Chebychev ya que estos no fueron tratados.

Llamaremos **esperanza matematica** (o valor esperado o media) a la funcion que, con la funcion de probabilidad (para una variable aleatoria continua o discreta) al numero que se obtiene de sumar cada valor a su probabilidad de suceso. Se denota con $E(X)$ o μ y se obtiene tal que:

$$E(X) = \sum_i x_i f(x_i) = \sum_i x_i P(X = x_i)$$

$$E(X) = \int_{-\infty}^{+\infty} x f(x)$$

En ocasiones, se puede pedir obtener la esperanza de una funcion definida para la variable aleatoria dada, se denotaria con $E(h(x))$ siendo $h(x)$ dicha funcion. Visto en el caso anterior, podemos decir que se habia obtenido con una $h(x) = x$, por lo tanto la transicion se da como sencilla de entender.

Es necesario saber que podemos denotar la esperanza para una funcion porque, para tratar con variables bidimensionales, es exactamente lo que haremos, teniendo entonces que, la esperanza para una variable bidimensional, seria:

$$E(h(X,Y)) = \sum_i \sum_j h(x_i, y_j) f(x_i, y_j) = \sum_i \sum_j x_i P(X = x_i, = y_j)$$

$$E(h(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x,y) f(x,y) dx dy$$

Distinguimos las distintas propiedades para la esperanza con la, veremos mas adelante, podremos definir la media muestral.

1. Dada una funcion, la esperanza para dicha funcion puede simplificarse extrayendo las constante. De esta forma, si tenemos $h(X) = aX + b$, donde a, b son constantes, podemos descomponer la obtencion de la esperanza a

   $$E(aX + b) = E(aX) + b = E(X)a + b$$

   teniendo unicamente que obtener la esperanza para la variable aleatoria y no para una funcion.

2. La esperanza de la suma de n variables es igual a la suma de las esperanzas de las n variables, tal que

   $$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

3. Si las n variables son independientes entre si, la propiedad 2 se cumple para el producto, teniendo que

   $$E(X_1 \cdot X_2 \cdot \ldots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \ldots \cdot E(X_n) \Leftrightarrow P(\bigcap_i X_i) = \prod_i P(X_i), P(X_i) \neq 0$$

Si tenemos definida la esperanza para una variable aleatoria, podemos definir la **varianza** de dicha variable como el valor que representa el grado de dispersion de los valores de la variable con respecto a su media (la esperanza). Se denota con σ^2 o $Var(X)$, teniendo valores siempre mayores o iguales a 0 y, es mas, la raiz cuadrada de la varianza es la **desviacion tipica**, denotada con σ. Veamos como se calcula:

$$Var(X) = E((X - E(X))^2) = E(X^2) - [E(X)]^2$$

Por comodidad, usar la segunda igualdad. Distingamos ahora las distintas propiedades:

1. Dada una funcion, la varianza para dicha funcion puede simplificarse extrayendo las contantes de ésta. De esta forma, si tenemos $h(X) = aX + b$, donde a, b son constantes, podemos descomponer la obtencion de la varianza a

   $$Var(aX + b) = Var(aX) = Var(X)a^2$$

   teniendo unicamente que obtener la varianza para la variable aleatoria y no para una funcion.

2. Dadas n variables independientes entre si, la varianza de la suma de éstas es igual a la suma de las varianzas de las n variables, tal que

$$Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$$

Es mas, si dado un par de variables donde se tiene $X - Y$ como la suma de estas, al aplicar la primera propiedad veremos que

$$Var(X-Y) = Var(X+(-1)Y) = Var(X)+(-1)^2Var(Y) = Var(X)+Var(Y) = Var(X+Y)$$

Ignoramos los momentos y lo relacionado a estos.

4.1. Covarianza y Correlacion

Al querer estudiar la relacion entre dos variables aleatorias, vamos a introducir la covarianza y la correlacion.

Veremos que cuando no existe ninguna relacion entre las variables, podremos afirmar que se tratan de variables independientes entre si.

Llamaremos **Covarianza** a la funcion

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

De esto, podremos afirmar las siguientes propiedades:

1. El valor obtenido puede ser positivo, negativo o nulo.
2. Si las variables son independientes, el valor tiene que ser nulo pero si obtenemos primero el valor y este es nulo, no significa que las variables sean nulas.
3. Podemos calcular la suma de dos variables aleatorias, sin importar que sean independientes (condicion necesaria en las propiedades indicadas previamente) si consideramos que la interseccion entre estas viene indicada en el valor de la covarianza multiplicada por 2, tal que
$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

El coeficiente de correlacion, denotado por ρ muestra el grado de relacion (lineal) entre las variables, se calcula con

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

El valor estara en el intervalo $[-1,1]$ y, si se da que tiene el valor de uno de los extremos, tendremos una realcion lineal perfecta ($Y = aX + b$). Cuando mas cerca del 0 este, menos intensa sera, siendo inexistente si el valor es nulo.

4.2. Esperanza condicional

Mirar si es complicado que la esperanza normal la calculabamos con la funcion de probabilidad que la esperanza condicional la calculamos con la funcion de probabilidad condicional... vamos a definirlas de todas formas.

Esperanza condicional de X dado $y = Y$:

$$E(X/y) = \sum_x x g_1(x/y), \text{ caso discreto}$$

$$E(X/y) = \int_{-\infty}^{+\infty} x g_1(x/y) dx, \text{ caso continuo}$$

Esperanza condicional de Y dado $x = X$:

$$E(Y/x) = \sum_y y g_2(y/x), \text{ caso discreto}$$

$$E(Y/x) = \int_{-\infty}^{+\infty} y g_2(y/x) dy, \text{ caso continuo}$$

4.3. Media muestral

Sabiendo que la esperanza es tambien la media, la media muestral para un muestra aleatoria simple $(X_1, \ldots, X_n)$ de la variable X donde $E(X) = \mu$ y $Var(X) = \sigma^2$, se definira como

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Las variables X_i son independientes entre si y todas tienen la misma distribucion que X, es decir, $E(X_i) = E(X)$ y $Var(X_i) = Var(X)$ (podriamos decir que es una variable aleatoria uniforme(?)), Vamos a ver como afectaria el obtener la esperanza y la varianza para la media muestral.

$$E(\overline{X_n}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n}[E(X_1) + \cdots + E(X_n)] = \frac{n\mu}{n} = \mu$$

$$Var(\overline{X_n}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \left(\frac{1}{n}\right)^2 [Var(X_1) + \cdots + Var(X_n)] = \frac{n\sigma^2}{n} = \frac{\sigma^2}{n}$$

Anotaciones mías

5. Distribuciones especiales

5.1. Distribucion de Bernoulli

Dada una variable aleatoria con solo dos posibles casos, a los cuales nos referiremos a ellos como caso exito (con una probabilidad de verificarse de p y comúnmente denotado con 1) y un caso fracaso (con una probabilidad de verificarse de $1 = 1 - p$ y comúnmente denotado con 0), llamaremos **distribucion de Bernoulli** a la distribucion de probabilidad para esta variable.

Considerar que en estos casos, la variable X no se ha usado para identificar el suceso de exito o fracaso, se usara para identificar el nº de pruebas realizadas o similar (en los ejercicios se tiene que definir).

5.2. Distribucion Binomial

Si tenemos n pruebas de Bernoulli independientes (para una variable aleatoria de solo dos casos, se hacen n pruebas para comprobar si se verifica el exito), tal que tendriamos $X = \sum_i^n X_i =$nº de existos en las n pruebas, definimos la **distribucion de probabilidad binomial** $B(n, p)$ donde la funcion de probabilidad asociada seria:

$$f(x) = P(X = x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x q^{n-x}, \ \ 0 \leq x \leq n$$

Recordamos que la operacion $\left(\begin{array}{c} n \\ x \end{array} \right) = \dfrac{n!}{x!(n-x)!}$

Con esta funcion, se obtiene la probabilidad de exito en la prueba x dada pero, por lo general, vamos a trabajar con los siguientes conceptos:

1. Obtencion de los parametros con las distintas pruebas de Bernoulli. Como sabemos que para todas las pruebas se va a tener la misma distribucion y que tendremos un total de n pruebas independientes, podemos aplicar propiedades de la Esperanza y la Varianza para descomponer valores u obtenerlos de forma mas sencilla, teniendo que:

$$E(X) = E(X_1 + \cdots + X_n) = \sum_i^n E(X_i) = \sum_i^n p = np$$
$$Var(X) = Var(X_1 + \cdots + X_n) = \sum_i^n Var(X_i) = \sum_i^n pq = npq$$

   Estas funciones obtienen estos valores ya que tenemos que considerar que la esperanza para un X_i dado seria

$$E(X_i) = 0{\cdot}q + 1{\cdot}p$$

   y para la varianza seria

$$Var(X_i) = (p) - (p^2) = (-(p-1))p = (-(1-q-1))p = qp$$

   esta ultima es un poco mas compleja, basicamente, sabiendo que tienes p veces p y le quitas una p al total, te quedan $p-1$, si obtienes el complementario, tendrias $p = 1 - q$ y simplificas la funcion. El desarrollo no es necesario, pero se deja por si acaso.

2. Si tenemos $X_1, \ldots, X_k$ que son variables aleatorias independientes ($P\left(\bigcap_i X_i\right) = \prod_i P(X_i)$) todas ellas con una distribucion binomial $B(n_i, p)$, entonces se tienes que $X = \sum_i^k X_i$ tiene una distribucion binomial $B(\sum_i^k n_i, p)$

Enfrentarse a un caso de Binomial

Habra que tener en cuenta que se usara no para obtener la probabilidad de una prueba, si no al valor acumulado de k pruebas en total (o de su complementario), y se usara la funcion de distribucion

$$F(k) = P(X = k) = \sum_i^k \left(\begin{array}{c} n \\ i \end{array} \right) p^i q^{n-i}$$

Por lo general, no habra que tener que calcularlo y bastara con revisar la tabla, fijandose en que definimos un conjunto de pruebas $X \sim B(n,p)$ y que queremos obtener el valor para $k \in X$, donde tendremos que buscar en la tabla para un n, k dados, la p del problema.

En ocasiones, se nos dara el valor mediante la Esperanza o la varianza y habra que despejar la probabilidad (conociendo el numero de pruebas).

Si se quiere obtener la probabilidad para uno de los casos en concreto, se puede obtener de la siguiente manera

$$P(X = k) = P(X \leq k) - P(X \leq (k-1)) = F(k) - F(k-1)$$

Unicamente tratar X con un numero de pruebas menor que 30 (o al menos que el acumulado que se desea obtener no sea mayor que 30)

5.3. Distribucion de Poission

Dejaremos las funciones y el procedimiento indicado pero veremos mas adelante que basta con obtener los parametros e interpretarlos adecuadamente en la tabla.

Dada una Binomial $B(n,p)$ la funcion de probabilidad que se obtiene al calcular el limite donde $n \mapsto \infty$ y $p \mapsto 0$ se llama distribución de Poisson, siendo esta:

$$lim_{\substack{n \mapsto \infty \\ p \mapsto 0}} f(k) = lim_{\substack{n \mapsto \infty \\ p \mapsto 0}} P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

Por lo general, nos referiremos a las distribuciones de Poisson con $P(\lambda)$, sabiendo que λ es la media para una X.

1. La obtencion de la media se puede obtener a traves de la Esperanza y/o la Varianza

$$E(X) = \lambda$$
$$Var(X) = E(X^2) - [E(X)]^2 = \lambda$$

   Por tu bien y por el mio, asumirlo... hay muchas pruebas con una probabilidad de exito muy baja.

2. Si tenemos $X_1, \ldots, X_k$ que son variables aleatorias independientes ($P\left(\bigcap_i X_i\right) = \prod_i P(X_i)$) todas ellas con una distribucion Poisson $P(\lambda_i)$, entonces se tienes que $X = \sum_i^k X_i$ tiene una distribucion de Poisson $P(\sum_i^k \lambda_i)$

Enfrentarse a un caso de Poisson

Normalmente usaremos Poisson cuando existe una $n > 30$ para Binomial, en este caso habra que considerar que, dado $X \sim B(n,p)$, tendremos $\lambda = np$, y por lo tanto, trataremos $P(np) = P(\lambda)$

En otros casos, se nos dira direcamente la media de sucesos para un numero de casos indeterminado. En este caso nos dan directamente el λ y lo que es mas probable que nos pregunte sera la probabilidad de que se den mas o menos casos. Considerar que la formula de la distribucion para Poisson es la siguiente:

$$F(k) = \sum_i^k e^{-\lambda}\frac{\lambda^i}{i!}$$

En la tabla de Poisson, tendremos definida $X \sim P(\lambda) : F(k) = P(X \leq k),\ k \in X$

5.4. Distribución Normal

Dada una variable aleatoria continua X con la siguiente funcion de denisdad

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

Se dira que tiene una distribucion Normal tipificada $X \sim N(0,1)$

La funcion de distribucion se denotara en este caso con $\Phi(x)$, donde:

$$\Phi(x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad -\infty < x < +\infty$$

$$\Phi(-x) = 1 - \Phi(x) \quad \Phi(x) = 1, x > 4 \quad \Phi(x) = 0, x < 4$$

Dada una variable aleatoria continua X con la siguiente funcion de denisdad

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \quad -\infty < y < +\infty$$

Se dira que tiene una distribucion de Normal con parametros μ y σ, que denotaremos como $Y \sim N(\mu, \sigma)$.

Recordamos que, en este caso, tendremos que

$$E(Y) = \lambda \text{ y } Var(Y) = E(Y^2) - [E(Y)]^2 = \sigma^2$$

Podemos relacionar la normal general con la tipificada, definiendo el valor tipificado en el proceso, para ello tenemos que:

1. Tener $X \sim N(0, 1)$ y $Y \sim N(\mu, \sigma)$, con $Y = \sigma X + \mu$
2. Al querer obtener $F(y)$, tendremos $P(Y \leq y)$, haremos una sustitucion de Y de tal manera que tendriamos

   $$P(\sigma X + \mu \leq y) = P\left(X \leq \frac{y-\mu}{\sigma}\right) = \Phi\left(\frac{y-\mu}{\sigma}\right)$$

   Donde distinguimos que el valor que se usa en la funcion de distribucion sera el que llamaremos valor tipificado.
3. Si queremos obtener el valor acumulado en un intervalo dado, se hace como siempre hemos hecho, restar la frecuencia del menor al mayor, pero usaremos los valores tipificados para ambos extremos.

Si tenemos $X_1, \ldots, X_k$ que son variables aleatorias independientes ($P\left(\bigcap_i X_i\right) = \prod_i P(X_i)$) todas ellas con una distribucion Normal cada una con λ_i y σ_i^2, entonces se tiene que $X = \sum_i^k X_i$ tiene una distribucion Normal con media $\lambda = \sum_i^k \lambda_i$ y varianza $\sigma^2 = \sum_i^k \sigma_i^2$

Dado el mismo contexto, habra que considerar que el valor tipificado de X se obtiene de la sigueinte operacion:

$$\frac{\sum_i^k X_i - \sum_i^k \mu_i}{\sqrt{\sum_i^k \sigma_i^2}}$$

La cual converge a una distribucion tipificada.

Enfrentarse a un caso de Normal

Pocas veces nos daran las cosas faciles para calcular la distribucion Normal, por lo general nos pediran obtenerla a traves de una Binomial o una de Poisson, para ello, habra que considerar las igualdades respecto a la Esperanza y la Varianza y a continuacion aplicarlos al valor tipificado para aproximar a la distribucion Normal tipificada.

- Dado $X \sim B(n,p)$, con $E(X) = np = \mu$ y $Var(X) = npq = \sigma^2$, podemos tipificar la Binomial, que se aproximara a $N(0,1)$, cuando n es suficientemente grande.

$$\frac{X - np}{\sqrt{npq}} \approx N(0,1)$$

- Dado $X \sim P(\lambda)$, con $E(X) = \lambda = \mu$ y $Var(X) = \lambda = \sigma^2$, podemos tipificar Poisson, que se aproximara a $N(0,1)$, cuando n es suficientemente grande.

$$\frac{X - \lambda}{\sqrt{\lambda}} \approx N(0,1)$$

6. Anexos

6.1. Repaso teoría de conjuntos

P.Asociativa	$\mapsto$	$(A \cup B) \cup C \Leftrightarrow A \cup (B \cup C)$
P.Conmutativa	$\mapsto$	$(A \cup B) \Leftrightarrow (B \cup A)$
P.Distributiva	$\mapsto$	$(A \cup B) \cap C \Leftrightarrow A \cap C \cup B \cap C)$
P.Transitiva	$\mapsto$	$(A \subset B) \wedge (B \subset C \Leftrightarrow A \subset C$
L.Morgan	$\mapsto$	$\overline{A \cup B} \Leftrightarrow \overline{A} \cap \overline{B} / \overline{A \cap B} \Leftrightarrow \overline{A} \cup \overline{B}$

Unión	$A \cup B := [x \vert x \in A \vee x \in B]$
Intersección	$A \cap B := [x \vert x \in A \wedge x \in B]$
Complementario	$\overline{A} := [x \vert x \notin A \vee x \in E]$
Resto	$A \setminus B := [x \vert x \in A \vee x \notin B] \Leftrightarrow A \cap \overline{B}$
Inclusión	$A \subset B := \forall x \in A \vert x \in B$

*E :=Conjunto referencia

6.2. Repaso de Integrales

En esta hoja de excel hay funciones que obtienen las primitivas (mirar aqui)

TABLA DE INTEGRALES INMEDIATAS

TIPOS	FORMAS	
	Simple	**Compuesta**
Potencial $n \neq -1$	$\int x^n dx = \frac{x^{n+1}}{n+1} + k$	$\int f'(x) f(x)^n dx = \frac{f(x)^{n+1}}{n+1} + k$
Logarítmico	$\int \frac{1}{x} dx = \int x^{-1} dx = \ln\lvert x\rvert + k$	$\int \frac{f'(x)}{f(x)} dx = \ln\lvert f(x)\rvert + k$
Exponencial	$\int e^x dx = e^x + k$ $\int a^x dx = \frac{a^x}{\ln a} + k$	$\int e^{f(x)} f'(x) dx = e^{f(x)} + k$ $\int a^{f(x)} f'(x) dx = \frac{a^{f(x)}}{\ln a} + k$
Seno	$\int \cos x dx = \operatorname{sen} x + k$	$\int \cos f(x) \cdot f'(x) dx = \operatorname{sen} f(x) + k$
Coseno	$\int \operatorname{sen} x dx = -\cos x + k$	$\int \operatorname{sen} f(x) \cdot f'(x) dx = -\cos f(x) + k$
Tangente	$\int \sec^2 x dx = \operatorname{tg} x + k$ $\int (1 + \operatorname{tg}^2 x) dx = \operatorname{tg} x + k$ $\int \frac{1}{\cos^2 x} dx = \operatorname{tg} x + k$	$\int \sec^2 f(x) \cdot f'(x) dx = \operatorname{tg} f(x) + k$ $\int \frac{f'(x)}{\cos^2 f(x)} dx = \operatorname{tg} f(x) + k$
Cotangente	$\int \operatorname{cosec}^2 x dx = -\operatorname{cotg} x + k$ $\int (1 + \operatorname{cosec}^2 x) dx = -\operatorname{cotg} x + k$ $\int \frac{1}{\operatorname{sen}^2 x} dx = -\operatorname{cotg} x + k$	$\int \operatorname{cosec}^2 f(x) \cdot f'(x) dx = -\operatorname{cotg} f(x) + k$ $\int (1 + \operatorname{cosec}^2 f(x)) f'(x) dx = -\operatorname{cotg} f(x) + k$ $\int \frac{f'(x)}{\operatorname{sen}^2 f(x)} dx = -\operatorname{cotg} f(x) + k$
Arco seno	$\int \frac{1}{\sqrt{1-x^2}} dx = \operatorname{arcsen} x + k$	$\int \frac{f'(x)}{\sqrt{1-f(x)^2}} dx = \operatorname{arcsen} f(x) + k$
Arco tangente	$\int \frac{1}{1+x^2} dx = \operatorname{arctg} x + k$ $\int \frac{1}{a^2+x^2} dx = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + k$	$\int \frac{f'(x)}{1+f(x)^2} dx = \operatorname{arctg} f(x) + k$ $\int \frac{f'(x)}{a^2+f(x)^2} dx = \frac{1}{a} \operatorname{arctg} \frac{f(x)}{a} + k$
Neperiano – Arco tangente	$\int \frac{Mx+N}{ax^2+bx+c} dx = \text{neperiano} + \text{arco tangente} + k$ $M \neq 0, \quad ax^2+bx+c$ irreducible	

Integral	Cambio recomendable
$\int (a^x) dx$	$z = a^x$
$\int (e^x)\, dx$	$z = e^x$
$\int (x, \ln x)\, dx$	$z = \ln x$
$\int (x, arc \ldots)\, dx$	$z = arc \ldots$
$\int (\sin^m x, \cos^n x)$	$z = \cos x$ si m impar $z = \sin x$ si n impar $z = \tan x$ si m, n pares
$\int (x, \sqrt{a^2 - b^2x^2}) dx$	$x = \frac{a}{b} \sin t$
$\int (x, \sqrt{a^2 + b^2x^2}) dx$	$x = \frac{a}{b} \tan t$
$\int (x, \sqrt{b^2x^2 - a^2}\,) dx$	$x = \frac{a}{b} \sec t$

A la hora de tratar con integrales dobles o parciales, aplicar la regla de prioridades (poneros parentesis si lo veis mas comodo) de tal manera que la variable que se este tratando en la parte mas interna ira primero, siendo la otra tratada como una constante y a continuacion, tratar a esta como la variable y la anterior como constante. De esta manera, si tenemos

$$\int \int (x + y) dx dy$$

primero integramos x porque es la mas interna y a continuacion y.

Hay que tener en cuenta que, en ocasiones, los valores de las limites de integracion vienen dados por funciones o estan en funcion de la otra variable, se esperan poner ejemplo a continuacion pero si no los ves puestos sera porque me da pereza hacerlo ahora y estoy con otras cosas.

Recordad despejar las contantes de una funcion hasta obtener la primitiva simple o, si no es posible, encontrar la formad e hacerla compuesta.

Podeis mirar esta pagina donde os resuelven las integrales por pasos, pulsar aqui. Ademas, si os acostumbrais a usar las distitnas opciones, sereis capaces de hacer integrales de dos variables con intervalos que dependen de otras variables.