TEMA II:
MODELOS DE DISTRIBUCIÓN DE PROBABILIDAD.

2.- **Variables aleatorias.**

2.1.- Variables aleatorias discretas.
- .-Función de probabilidad o función de cuantía.
- .- Función de distribución.

2.2.- Variables aleatorias continuas.
- .- Función de densidad de probabilidad.
- .- Función de distribución.

2.3.- Medidas características de una variable aleatoria.
- .- Media o esperanza matemática y varianza.
- .- Desigualdad de Tchebyshev.
- .- Momentos.
- .- Función generatriz de momentos

2.4.- Algunos ejemplos de distribuciones discretas.
- .- Distribución Binomial
- .- Distribución de Poisson.
- .- Otras distribuciones discretas: Geométrica, Binomial Negativa, Hipergeométrica.

2.5.- Algunos ejemplos de distribuciones continuas.
- .- Distribución Exponencial.
- .- Distribución Normal.

2.6.- La Normal como aproximación de otros distribuciones.
- .- Teorema Central del Límite.
- .- Relación entre la Binomial, la Poisson y la Normal.

2.7.- Distribuciones asociadas a la distribución Normal
- .- c^{2} de Pearson.
- .- F de Fisher.
- .- t de Student.

2.8.- Otras distribuciones continuas : Gamma, Beta, Weibull etc.

DISTRIBUCIONES DE PROBABILIDAD.

Segunda parte.

1. - Se lanza al aire cinco veces una moneda y se mira el número de apariciones de la cara. Escribir la ley de distribución del número de veces que aparece cara y obtener la probabilidad de:

i) Obtener al menos una cara. i) Obtener mas de tres caras.

2.- El cuestionario de unas oposiciones consta de 100 temas entre los que el opositor ha de contestar uno elegido al azar. Supongamos que un opositor conoce sólo 80 temas. Repite la oposición en idénticas condiciones de preparación 5 veces. Hallar:

i) Probabilidad de que las cinco veces le hayan salido temas desconocidos.
ii) Probabilidad de que apruebe la quinta vez.

3.- Al inspeccionar 1500 tornillos producidos por una máquina automática se han encontrado 150 defectuosos. Se colocan 5 tornillos en un aparato. ¿Cual es la probabilidad de que r de ellos (R = 0, 1, 2, 3, 4, 5) sean defectuosos?. Calcular la esperanza y la varianza.

4.- Una editorial edita 50000 ejemplares de un mismo libro y se estima que la probabilidad de que un libro presente defectos en su encuadernación es 0.0001. Hallar la probabilidad de que:

i) La tirada no contenga ningún libro mal encuadernado.
ii) La tirada contenga tres libros mal encuadernados.

5.- Una máquina dedicada a la impresión de tarjetas de visita realiza en un día 20000. La probabilidad de que una sea defectuosa es 0.0001. Se pide:

i) La función de cuantía.
ii) La probabilidad de obtener todas las tarjetas perfectas.
iii) La probabilidad de obtener dos tarjetas defectuosas.
iv) La probabilidad de obtener por lo menos dos tarjetas defectuosas.

6.- Una máquina automática de imprenta está compuesta por 500 interruptores independientes uno de otro. La probabilidad de fallo de cualquiera de ellos durante un periodo de tiempo t es 0.001. Determinar:

i) La probabilidad de que durante el periodo de tiempo t fallen tres interruptores.
ii) " " " " mas de tres.

iii) " " " " a lo sumo tres.

7.- Una máquina de fabricar caramelos tiene r roturas con probabilidad.

$$P(X=r)=\frac{k3^{-r}}{r!}$$

Calcular el valor de la constante y la probabilidad de que se produzca al menos una rotura.

8.- El número de errores tipográficos por página cometidos en la impresión de un libro de 95 páginas tiene la siguiente distribución de frecuencias:

Nº de errores	Frecuencia
0	39
1	31
2	15
3	7
4	2
5	1

Admitiendo que la variable errores cometidos sigue una distribución de Poisson, determinar el parámetro y obtener las frecuencias esperadas a partir del ajuste por la distribución de Poisson.

9.- El número de llamadas telefónicas recibidas en una central telefónica en un intervalo de tiempo se ha tabulado arrojando los siguientes resultados. Ajustar una distribución de Poisson a estos datos.

Nº de llamadas	frecuencia
0	144
1	136

2	74
3	32
4	11
5	2
6	1

10.- La probabilidad de que se rompa la lámpara de un televisor en un mes es del 2%. Si tenemos 5 años el televisor, calcular:

i) La ley de probabilidad del nº de roturas.
ii) Media y desviación típica.
iii) Probabilidad de que no haya ninguna rotura.
iv) Probabilidad de que haya mas de una rotura.

11.- Una máquina de empaquetado de cigarrillos realiza por término medio un 1% de defectuosos. Los paquetes se guardan en cajas de 200. Se pide:

i) ¿Cuántas cajas no tienen ningún paquete defectuoso?
ii) ¿Cuántas cajas contienen un paquete defectuoso?
iii) ¿ " " a lo sumo un paquete defectuoso?

12.- En una distribución N(5, 2) determinar las siguientes probabilidades:

a) $P(7\leq X\leq 10)$ b) $P(-1\leq X\leq 10)$ c) $P(-2\leq X\leq -1)$

13.- En una distribución N(3, 2) obtener los puntos a y b que verifican:

i)$P(X\leq a)=0.8413$ ii)$P(X\leq b)=0.9772$

14.- En una distribución N(0, 1) encontrar :

i) $P(|X|\leq a)$ ii) $P(X\geq a)$ iii) $P(|X|\geq a)$

Aplicar el resultado cuando a=0.5 y cuando a=1.23

15.- Una variable aleatoria se distribuye normalmente con media 20. La probabilidad de que pertenezca al intervalo (25, 30) es del 25%. Calcular $P(10<X<15)$.

16.- Las puntuaciones obtenidas por los alumnos de una facultad en la asignatura de estadística siguen una distribución normal. Determinar la media y la desviación típica sabiendo que:

i) $P(X\leq 3)=0.1587$ ii) $P(X>9)=0.0228$

17.- Una máquina automática fabrica tornillos. La longitud del tornillo es una variable aleatoria normal de media 5cm. Y se sabe que toma valores en el intervalo (3.2, 6.8). Determinar la probabilidad de que un tornillo elegido al azar tenga una longitud:

i) Mayor que 5.2 cm. ii) Menor que 4.4 cm.

18.- Una central telefónica puede establecer 670 conexiones por minuto. El número de conexiones solicitadas por minuto es una v.a. de Poisson de parámetro λ=625. Determinar la probabilidad de que en un minuto determinado la central esté saturada de llamadas.

19.- Se lanza al aire una moneda 400 veces y sea X la variable que representa el número de veces que sale cara. Se pide:

a)Ley de probabilidad de la variable. b)Esperanza matemática y varianza. c)$P(X>220)$ d)$P(175\leq X\leq 230)$

20.- Un dado corriente se lanza al aire 720 veces. Sea X la variable que representa el número de veces que sale el 6. Se pide:

a)La distribución de X. Media y varianza de X. b) $P(100\leq X\leq 125)$ $P(X>150)$

21.- Sabiendo que la probabilidad de hacer blanco en una diana es 0.01,¿cuantos disparos deberán realizarse para, con una probabilidad del 95%, alcanzar el objetivo por lo menos una vez?.

22.- Si las llamadas telefónicas en una centralita siguen una distribución de Poisson de parámetro λ=3 llamadas cada 5 minutos, calcular la probabilidad de:

a) Seis llamadas en cinco minutos. b) Tres en diez.
c) Mas de quince en un cuarto de hora. d) Dos en un minuto.
e) Calcular la probabilidad de que transcurran cinco minutos sin ninguna llamada.

23.- Supongamos un experimento que tiene una probabilidad de éxito igual a 0.01.Calcular cuantas veces debe repetirse para que la probabilidad de al menos tres éxitos sea como mínimo de 0.9.

24.-Supongamos una calculadora que tiene cuatro circuitos impresos Sea p_i la probabilidad de que una calculadora que ha sido enviada para ser reparada necesite i circuitos nuevos. Se conoce que $p_1=1/2$, $p_2=1/4$, $p_3=p_4=1/8$. Si se envían 10000 calculadoras a reparar al año, cual es la probabilidad de necesitar mas de 18875 circuitos?

25.-Una empresa recibe piezas de un proveedor en lotes de 2000 que se someten al siguiente control de calidad: se toman 20 piezas al azar y si hay mas de una defectuosa se rechaza el lote; en otro caso se acepta. La calidad garantizada por el proveedor es de un 8 por mil de defectuosas. Calcular la probabilidad de :

- i) Aceptar un lote que contenga un 2% de defectuosas.
- ii) Rechazar un lote que debería haber sido aceptado por contener sólo el 8 por mil de defectuosas.

26.- La media de las calificaciones obtenidas en un test de aptitud por los alumnos de un centro fue 400 y la desviación típica 100. Si las calificaciones siguen una distribución normal, calcular:

- i) ¿ Qué porcentaje de alumnos tuvieron calificación superior a 500 puntos?
- ii) ¿ Cual es la probabilidad de que elegido un alumno al azar, su calificación difiera de la media en menos de 150 puntos?

27.- Las puntuaciones obtenidas por los 300 niños de un colegio al aplicarles un test de aptitud numérica , sigue una distribución normal de media 24 y desviación típica 4. Calcular:

- i) El cuartil inferior.
- ii) El cuartil superior.
- iii) ¿Cuál es la probabilidad de obtener puntuación igual o inferior a 16?.
- iv) ¿Cuántos niños tienen una puntuación igual o mayor que 26?

28.- Una variable aleatoria se distribuye uniformemente en el intervalo [a,2a]. Calcular la función de densidad, la función de distribución, la media y la varianza.

29. - Hallar los valores críticos χ^2_α de χ^2 en los siguientes casos.

- i) $\alpha=0.05$ y n=15.
- ii) $\alpha=0.90$ y n=22
- iii) $\alpha=0.02$ y n=10

30. - Calcular los valores de t_α correspondientes a una t de Student en los siguientes casos.

- i) El área a la derecha de t_α es 0.30 y n=15.
- ii) El área a la izquierda de t_α es 0.10 y n=12.
- iii) El área a la izquierda de t_α es 0.20 y n=38.

31. - Calcular los valores de F_α correspondientes a una distribución F de Snédecor en los casos siguientes.

- i) $\alpha=0.01$ y (4,15) grados de libertad.
- ii) $\alpha=0.05$ y (12,2) grados de libertad.

Ejercicios de exámenes del curso 1999-2000 que corresponden a Probabilidad y a Distribuciones.

Junio

.- Una urna A contiene 4 bolas negras y 6 rojas, y una urna B contiene 5 bolas negras y 5 rojas. Se lanza una moneda. Si sale cara se ponen dos bolas de A en B y si sale cruz se pone una de A en B. A continuación sacamos una bola de B y resulta ser roja. Calcular la probabilidad de que en la moneda saliera cara.

.- Si X sigue una distribución Normal con media 10 y desviación típica 2, se pide:

a) $P(X\leq 12)$
b) $P(|X|\leq 13)$
c) $P(|X-10|\leq 1)$
d) Hallar el percentil 0,95 de la distribución.
e) Hallar el valor de a para que $P(|X-10|\leq a)=0,95$

Septiembre.

.- Supongamos un sistema con 9 componentes que requiere para su funcionamiento que al menos 6 funcionen. Si la probabilidad de funcionamiento de una componente es 0,95, calcular la probabilidad de que el sistema funcione.

Si el sistema tiene 100 componentes y funciona si hay 60 disponibles, ¿cuál es la probabilidad de que funcione?.

.- Una compañía de seguros de automóviles clasifica a los conductores en tres clases:

A: alto riesgo. B: riesgo medio. C: bajo riesgo.

La clase A constituye el 30% de los conductores que suscribe un seguro con la compañía y la probabilidad de que uno de esos conductores tenga un accidente en un año es 0,1. Los datos correspondientes para la clase B son 50% y 0,03, y para C 20% y 0,01.

a) Calcular la probabilidad de que un conductor asegurado en la compañía sufra un accidente en un año.
b) Un determinado cliente contrata una póliza y tiene en el primer año un accidente. ¿Cuál es la probabilidad de que sea de la clase A?
c) Si un asegurado lleva diez años sin sufrir accidentes y suponemos que los años son independientes en cuanto al número de accidentes, ¿cuál es la probabilidad de que sea de la clase C?

Curso 2001/2002

1.- Una editorial sabe por la experiencia acumulada durante años que un 5% de los libros que edita tiene una encuadernación imperfecta. Empaqueta los libros en paquetes de 30 libros y después estos paquetes los pone para su transporte en cajas de madera con 50 paquetes cada una . Antes de enviarlos a los distribuidores hace el siguiente control. Elige una caja al azar y de esta un paquete cualquiera del que toma 7 libros y revisa su encuadernado.Si encuentra al menos uno defectuoso entonces revisa todos los del paquete y si encuentra más de 3 defectuosos revisa todos los de la caja . Se pide:

a) Calcular la probabilidad de que tenga que revisar todo el paquete.
b) Calcular la probabilidad de que tenga que revisar toda la caja.
c) Si la editorial está en lo cierto en cuanto al porcentaje de defectuosos,¿ cual es la probabilidad de que la caja tenga más de 50 libros defectuosos?.

2.- Calcular las siguientes probabilidades :

a) $P(2 \leq x \leq 7)$ si x tiene una distribución B (9 , 0.7)
b) $P (3 < x \leq 8)$ si x sigue una distribución de Poisson de parámetro 3.
c) $P (|x-10| < 3)$ si x sigue una distribución N (8 , 6)
d) Calcular n para que $P (t \leq t_{n\,,\,0.975}) = 2.16$. t es una (t – de Student).
e) Calcular α para que $P (\chi^2 \geq \chi^2_{\alpha\,,16}) = 10.312$.

2.- Variables aleatorias.

Se denomina variable aleatoria a una variable numérica cuyos valores vienen determinados por el azar, es decir, asigna un valor numérico a cada uno de los sucesos que resultan al realizar un experimento aleatorio.

Ejemplos.

Experimento aleatorio	*Variable aleatoria*
Lanzar un dado.	X = " Número que figura en la cara del dado" $X\in \{1, 2, 3, 4, 5, 6\}$
Lanzar un dardo.	(X, Y) = " Posición en la diana". $(X, Y) \in \{(x, y) \in \mathbf{R}^2 / x^2 + y^2 \leq r^2\}$
Romper en dos trozos una varilla.	X = " Longitud del trozo mas largo" $X\in (0, L)$
Sexo de un recién nacido.	X = " Niño o Niña". $X\in \{0, 1\}$

La definición permite trabajar considerando como un suceso el hecho de que la variable tome un determinado valor "$X = x_i$" o pertenezca a un intervalo "$X \geq x_i$" o " $x_i \leq X \leq x_j$" . De cada suceso de este tipo se puede determinar su probabilidad , es decir $P(X = x_i)$, $P(X \geq x_i)$ y $P(x_i \leq X \leq x_j)$. Se dice que se ha definido un ***modelo de distribución de probabilidad*** para una variable cuando se tienen los valores que toma la variable y la probabilidad de que los tome.

Según sea el conjunto numérico en el que toman valores las variables aleatorias pueden ser ***discretas o continuas***.

Función de distribución.

Se llama función de distribución de una variable X a la función $F(x) = P(X \leq x)$ $\forall x\in R$.

Propiedades de la función de distribución.

i) $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$.

ii) $0 \leq F(x) \leq 1$.

iii) F es una función monótona no decreciente, es decir, si $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$.

iv) F se define en todo el eje real y se conviene que $\lim_{x\to+\infty} F(x) = 1$ y $\lim_{x\to-\infty} F(x) = 0$.

v) F es continua por la derecha, es decir:

$$\lim_{x\to a^+} F(x) = F(a) \qquad \text{y} \qquad \lim_{x\to a^-} F(x) = F(a) - P(X = a)$$

Demostración.

i) $(X \le x_2) = (x_1 < X \le x_2) \cup (X \le x_1)$, y como los sucesos son incompatibles $P(X \le x_2) = P(x_1 < X \le x_2) + P(X \le x_1)$.

iii) $F(x_2) - F(x_1) = P(x_1 < X \le x_2) \ge 0$.

v) Si $x > a \Rightarrow P(a < X \le x) = F(x) - F(a)$

$$\Rightarrow \begin{cases} \lim\limits_{x\to a^+} F(x) = F(a) + \lim\limits_{x\to a^a} P(a < X \le x) \\ \bigcap\limits_{x\to a^+} (a, x] = F \Rightarrow \lim\limits_{x\to a^a} P(a < X \le x) = 0 \end{cases}$$

Si $x < a \Rightarrow P(x < X \le a) = F(a) - F(x) \Rightarrow F(x) = F(a) - P(x < X \le a)$

$$\Rightarrow \begin{cases} \lim F(x) = F(a) - \lim\limits_{x\to a^-} P(x < X \le a) \\ \bigcap\limits_{x\to a^-} (x, a] = \{a\} \Rightarrow \lim\limits_{x\to a^-} P(x < X \le a) = P(X = a) \end{cases}$$

Para que la función de distribución sea continua por la izquierda es necesario que $P(X = a) = 0\ \forall a \in R$.

2.1.- Variables aleatorias discretas.

Una variable aleatoria se dice discreta cuando toma valores en conjuntos discretos, bien sean finitos o infinitos numerables.

Función de probabilidad o función de cuantía.

Sea $X \in \{x_1, x_2, \ldots, x_n\}$(n no necesariamente finito) una variable aleatoria discreta. Se llama función de probabilidad o función de cuantía a la función que para cada valor de la variable xi, proporciona la $P(X = x_i)$. Presentamos la función en la siguiente tabla.

X	x_1	x_2	…	x_n
$P(X = x_i)$	$P(X = x_1)$	$P(X = x_2)$	…	$P(X = x_n)$

Observar que $\sum\limits_{i=1}^{n} P(X = x_i) = 1$

Función de distribución de una variable discreta.

La función de distribución de una variable discreta en los valores de la variable será:

$F(x_1) = P(X \le x_1) = P(X = x_1)$

$F(x_2) = P(X \le x_2) = P(X = x_1) + P(X = x_2)$.

……….

$F(x_n) = P(X \le x_n) = P(X = x_1) + P(X = x_2) + \ldots + P(X = x_n)$.

Por lo que F(x) vendrá dada por:

$$F(x) = \begin{cases} 0 & x < x_1 \\ P(X = x_1) & x_1 \leq x < x_2 \\ P(X = x_1) + P(X = x_2) & x_2 \leq x < x_3 \\ & \\ P(X = x_1) + P(X = x_2) + \ldots + P(X = x_j) & x_j \leq x < x_{j+1} \\ & \\ 1 & x \geq x_n \end{cases}$$

Propiedades de la función de distribución de una v.a. discreta.

La función de distribución en el caso discreto tiene además las siguientes propiedades.

i) $P(a \leq X \leq b) = F(b) - F(a) + P(X = a)$.

ii) $P(a < X < b) = F(b) - F(a) - P(X = b)$.

iii) $P(a \leq X < b) = F(b) - F(a) - P(X = b) + P(X = a)$.

Demostración.

$(X \leq b) = (X \leq a) \cup (a < X \leq b)$

i) $(a \leq X \leq b) = (X=a) \cup (a < X \leq b)$.

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

ii) $(a < X \leq b) = (X=b) \cup (a < X < b)$.

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

iii) $(a \leq X < b) = (X=a) \cup (a < X < b)$.

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

EJEMPLO I.

En el lanzamiento de un dado escribir las funciones de cuantía y de distribución de probabilidad de la variable número de puntos en la cara superior.

Solución.

Función de cuantía.

X	1	2	3	4	5	6
$P(X = x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

Función de distribución.

$F(1) = P(X \leq 1) = P(X = 1) = 1/6$

$F(2) = P(X \leq 2) = P(X = 1) + P(X=2) = 2/6$.

……….

$F(6) = P(X \leq 6) = P(X = 1) + P(X = 2) + \ldots + P(X = 6) = 1$

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \leq x < 2 \\ 2/6 & 2 \leq x < 3 \\ 3/6 & 3 \leq x < 4 \\ 4/6 & 4 \leq x < 5 \\ 5/6 & 5 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

EJEMPLO II.

Se lanza tres veces una moneda. Escribir la función de cuantía y la función de distribución de la variable número de caras.

Solución.

Función de cuantía.

X	3	2	1	0
$P(X = x_i)$	1/8	3/8	3/8	1/8

Función de distribución.

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 3/8 & 1 \leq x < 2 \\ 3/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

2.2.- Variables aleatorias continuas.

Se dice que una v.a. es continua cuando es continua su función de distribución. Esto significa que :

$$\lim_{x \to a^+} F(x) = \lim_{x \to a^-} F(x) = F(a) \Rightarrow \begin{cases} P(X = a) = 0 \\ P(a < X \leq x) = F(x) - F(a) \to 0 \quad \text{cuando } x \to a \end{cases}$$

La continuidad implica que es cero la probabilidad de un punto y que intervalos pequeños tienen un probabilidad pequeña. Son variables continuas las que toman valores en un intervalo.

Función de densidad de probabilidad.

Se dice que una función real de variable real f(x) es una función de densidad de una v.a. X si verifica las siguientes propiedades.

Propiedades de la función de densidad.

i) $f(x) \geq 0 \ \forall x \in R$.

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

iii) $P (x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)\, dx$

Como consecuencia inmediata de la definición se tiene que
$P(x_1 < X < x_2) = P(x_1 \le X \le x_2) = P(x_1 \le X < x_2) = P(x_1 < X \le x_2)$.

Función de distribución.

Si X es una v.a. continua con función de densidad de probabilidad f(x) su función de distribución F(x) vendrá dada por:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt \quad \forall x \in R.$$

Como consecuencia inmediata de esta definición tenemos que :

i) $P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)\ dx = F(x_2) - F(x_1)$

ii) $F'(x) = f(x)$.

Es conveniente resaltar tres aspectos .

i) La función de densidad desempeña en el caso continuo el mismo papel que la función de cuantía en el caso discreto.

ii) En el caso continuo la integral sustituye a la suma.

iii) Cuando la variable es continua la probabilidad de que tome un valor concreto es cero.

EJEMPLO I.

Se dice que una variable aleatoria tiene una distribución uniforme en un intervalo (a, b) si su función de densidad es constante en dicho intervalo y nula fuera de él.

$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in (a,b) \\ 0 & x \notin (a,b) \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{a}^{x} \frac{1}{b\text{ - }a}\ dt = \frac{1}{b-a}(x-a) \qquad a < x < b$$

Según esto la función de distribución es:

$$F(x) = \begin{cases} 0 & x \le a \\ \dfrac{1}{b-a}(x-a) & a < x < b \\ 1 & x \ge b \end{cases}$$

EJEMPLO II.

Se dice que una variable tiene una distribución exponencial de parámetro λ si su función de densidad es:

$$f(x) = \begin{cases} ?e^{-?x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

y su función de distribución es $F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-?x} & x \geq 0 \end{cases}$.

2.3.- Medidas características de una variable aleatoria.
Media o esperanza matemática .

La media o esperanza matemática de una variable aleatoria es una medida del comportamiento central de la variable. Se representa $E[X] = \mu$ y se define de manera diferente según que la variable sea discreta o continua.

Caso discreto.

$$E[X] = \sum_{i=1}^{n} x_i \, P(X = x_i)$$

Si la variable toma valores en un conjunto numerable esa suma sería una serie.

Caso continuo.

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

Propiedades de la esperanza.

i) Si k es una constante $E[k] = k$.

ii) $E[a + b\,X] = a + b\,E[X]$

iii) Si $j(x)$ es una función de la variable aleatoria X.

$$E[\, j(X)\,] = \sum_{i=1}^{n} j(x_i)\, P(X = x_i) \text{ en el caso discreto.}$$

$$E[\, j(X)\,] = \int_{-\infty}^{+\infty} j(x) \; f(x) \; dx \text{ en el caso continuo.}$$

iv) Si X, Y, ... , W son variables aleatorias definidas sobre un mismo espacio muestral $E[X + Y + ... + W] = E[X] + E[Y] + ... + E[W]$.

EJEMPLO I.

Calcular la esperanza de los beneficios obtenidos por una compañía de seguros al hacer un seguro cuya prima anual es r, la probabilidad de siniestro es p y la cantidad asegurada es M.

Solución.

La variable aleatoria X representa el beneficio. Su función de cuantía es:

X	r	-(M-r)
$P(X = x_i)$	1-p	p

$E[X] = r(1-p)-(M-r)p = -Mp + r$

Si $r > Mp$ el beneficio estará asegurado.

EJEMPLO II.

En la ruleta hay 37 números que van desde el cero hasta el 36. De ellos 18 son rojos y 19 negros. Calcular el beneficio esperado con una apuesta de 1000 pts:

i) A un número cualquiera. Si acierta gana 35000pts.
ii) A rojo contra negro. Si acierta gana 1000.

Solución.
Llamamos X a la variable beneficio obtenido.
Función de cuantía en el caso i)

X	35000	-1000
$P(X = x_i)$	1/37	36/37

E[X] = 35000x1/37-1000x36/37= -27

En el segundo caso la función de cuantía será.

X	1000	-1000
$P(X = x_i)$	18/37	19/37

E[X] = 1000x18/37 - 1000x19/37 = -27

Las dos apuestas tienen la misma esperanza de pérdida.

EJEMPLO III. La esperanza de la distribución uniforme es

$$E[X] = \int_{-\infty}^{+\infty} x \frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{b+a}{2}$$

EJEMPLO IV. La esperanza de la distribución exponencial es

$$E[X] = \int_{-\infty}^{+\infty} x\, f(x)\,dx = \int_0^{+\infty} x \quad l \quad e^{-l\ x}\,dx = \frac{1}{l}$$

El resultado se ha obtenido haciendo por partes la integral impropia.

Varianza.

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

Caso discreto.

$$Var[X] = \sum_{i=1}^{n} (x_i - E[x])^2\, P(X = x_i)$$

Caso continuo.

$$Var[X] = \int_{-\infty}^{+\infty} (x - E[x])^2\, f(x)\,dx$$

Propiedades de la varianza.

i) $Var[a+bX] = b^2\, Var[X]$.
ii) Si X es una v.a. tal que $E[X] = \mu$ y $Var[X] = \sigma^2 \Rightarrow X^* = \frac{X-\mu}{s^2}$ es tal que $E[X^*] = 0$ y $Var[X^*] = 1$. X* se denomina variable tipificada.
iii) Desigualdad de Tchebyshev.

$$P(|X-\mu| \geq e) \leq \frac{Var[x]}{e^2}.$$

Momentos.

Momento de orden k respecto al origen.

$$m_k = E[X^k] = \begin{cases} \sum_{i=1}^{n} x_i^k P(X = x_i) \\ \int_{-\infty}^{\infty} x^k f(x)dx \end{cases}$$

Momento de orden k respecto de la media.

$$\mu_k = E[(X - E[x])^2] = \begin{cases} \sum_{i=1}^{n} (x_i - m)^k P(X = x_i) \\ \int_{-\infty}^{\infty} (x - m)^k f(x)dx \end{cases}$$

Coeficiente de Asimetría.

$$CA = \frac{\mu_3}{s^3}$$

Coeficiente de Apuntamiento.

$$CA_p = \frac{\mu_4}{s^4}$$

Función generatriz de momentos

$$?(t) = E[e^{tx}] = \begin{cases} \sum_{i=1}^{n} e^{tx_i} P(X = x_i) \\ \int_{-\infty}^{\infty} e^{tx} f(x)dx \end{cases}$$

Propiedades de la función generatriz de momentos.

i) $\psi(0) = 1$

ii) La función generatriz de momentos se llama así porque es una herramienta importante en el cálculo de momentos y esto es fundamental para el cálculo de la esperanza y la varianza de las distintas distribuciones.

$$?(t) = E[e^{tx}] = \begin{cases} \sum_{i=1}^{n} e^{tx_i} P(X = x_i) \\ \int_{-\infty}^{\infty} e^{tx} f(x)dx \end{cases}$$

$$?'(t) = \begin{cases} \sum_{i=1}^{n} x_i \, e^{tx_i} \, P(X = x_i) \\ \int_{-\infty}^{\infty} x \, e^{tx} \, f(x) \, dx \end{cases} = E[X \, e^{tx}]$$

$$?''(t) = \begin{cases} \sum_{i=1}^{n} x_i^2 \, e^{tx_i} \, P(X = x_i) \\ \int_{-\infty}^{\infty} x^2 \, e^{tx} \, f(x) \, dx \end{cases} = E[X^2 \, e^{tx}]$$

....

$$?^k(t) = \begin{cases} \sum_{i=1}^{n} x_i^k \, e^{tx_i} \, P(X = x_i) \\ \int_{-\infty}^{\infty} x^k \, e^{tx} \, f(x) \, dx \end{cases} = E[X^k \, e^{tx}]$$

y $?^k(0) = E\ [X^k]$ con lo que $?'(0) = E\,[X]$, $?''(0) = E\,[X^2]$

Var $[X] = ?''(0) - (?'(0))^2$.

2.4.- Algunos ejemplos de distribuciones discretas.

- Distribución de Bernouilli.

Se denomina proceso de Bernouilli a un experimento aleatorio que consiste en observar elementos de una población bajo las siguientes condiciones.

i) Las observaciones se clasifican en dos categorías que se denominan éxitos y fracasos.

ii) La proporción de éxitos y fracasos es constante y no se modifica cualquiera que sea la cantidad observada. Esto implica que las observaciones se hacen con reemplazamiento . Llamaremos p a la probabilidad de éxito y q = 1-p a la de fracaso.

iii) Las observaciones son sucesos independientes.

Ejemplos de este proceso son observar el sexo de un recién nacido, observar la aparición de un elemento defectuoso en un proceso de fabricación, etc. Existen diferentes variables aleatorias que pueden definirse en un proceso de Bernouilli cada una de ellas con distinta distribución de probabilidad.

En el proceso de Bernouilli, se define una v.a. que toma el valor 1 si el resultado de la prueba ha sido un éxito y el valor 0 si el resultado ha sido un fracaso. Según esto P(X = 1) = p y P(X = 0) =1-p.

E[X] = 1 P(X = 1) + 0 P(X = 0) = p

Var $[X] = (0-p)^2$ P(X = 0) + $(1-p)^2$ P (X = 1) = p(1-p) = pq.

Des [X] = $\sqrt{pq}$

- Distribución Binomial.

Se llama v.a. Binomial a la que describe el número de éxitos obtenidos cuando se realizan n pruebas de Bernouilli. $X \in \{0, 1, 2, ..., n\}$. El suceso X = k representa que se han obtenido k éxitos, y por tanto $n-k$ fracasos, al realizar n pruebas de Bernouilli. La probabilidad de este resultado teniendo en cuenta que las pruebas de Bernouilli son sucesos independientes es $p^k(1-p)^{n-k}$ y como con k éxitos habrá $\binom{n}{k}$ resultados diferentes se tendrá que $P(X = k) = \binom{n}{k} p^k(1-p)^{n-k}$ que es la función de cuantía de la distribución.

Observar que $\sum_{k=1}^{n} \cdot \binom{n}{k} p^k (1-p)^{n-k} = (p + q)^n = 1$.

La función de distribución de la binomial será

$$F(x) = \sum_{k \le x} \binom{n}{k} p^k (1-p)^{n-k}$$

Par indicar que una v.a. es binomial escribiremos $X \overset{D}{=} B(n, p)$ siendo n el número de pruebas de Bernouilli y p la probabilidad de éxito en cada prueba.

<u>Esperanza y Varianza de la distribución Binomial</u>.

Para calcular la esperanza y la varianza se usa la función generatriz de momentos.

$$?(t) = E\left[e^{tx}\right] = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} e^{tk} \binom{n}{k} (pe^t)^k (1-p)^{n-k} =$$

$$= (e^t p + q)^n$$

$$?'(t) = n(e^t p + q)^{n-1} pe^t \Rightarrow ?'(0) = np$$

$$?''(t) = n(n-1)(e^t p + q)^{n-2} p^2 e^{2t} + n(e^t p + q)^{n-1} pe^t \Rightarrow$$

$$?''(0) = n(n-1)(p+q)^{n-2} p^2 + n(p+q)^{n-1} p = n(n-1)p^2 + np$$

De donde se concluye que E [X] = n p y Var [X] = n p q.

EJEMPLO I.

La probabilidad de que un alumno apruebe un examen es 0.25. En un grupo de 4 estudiantes calcular:

i) Probabilidad de que los cuatro aprueben.
ii) Probabilidad de que no apruebe ninguno.
iii) Probabilidad de que sólo uno apruebe.
iv) Probabilidad de que al menos dos aprueben.
v) Probabilidad de que a lo sumo tres aprueben.

<u>Solución.</u>

X = “ número de alumnos de es grupo de cuatro que aprueban” $\overset{D}{=} B(4,0.25)$

i) P(X = 4)=0.0039

ii) P(X = 0)=0.3164
iii) P(X = 1)=0.4219
iv) P(X≥2)=1-P(X≤ 1)=1-0.7383
v) P(X≤3)=0.9961

EJEMPLO II.

Un sistema con nueve componentes requiere para su funcionamiento que al menos seis estén disponibles. Si la probabilidad de que una componente funcione es 0.95, calcular la probabilidad de que el sistema funcione.

Solución.

Sea Y = "número de componentes que no funcionan en el sistema "$\stackrel{D}{=}$B(9,0.05) P(Y≤3)=0.9994.

Distribución de Poisson.

Consideremos un experimento aleatorio en el que observamos la aparición de sucesos puntuales sobre un soporte continuo, por ejemplo:

i) Estudio de las averías de máquinas en un periodo de tiempo.
ii) Llegadas de aviones a un aeropuerto en un intervalo temporal.
iii) Llamadas recibidas en una central telefónica en un intervalo temporal.
iv) Corrosión en puntos de la superficie de una tubería.

Condiciones del proceso de Poisson.

a) El proceso es estable, es decir, produce a lo largo del intervalo soporte (espacial, temporal, etc.) un número medio de sucesos que es constante y que se representa por un parámetro λ.
b) Los sucesos ocurren de forma independiente.

Una variable aleatoria discreta se dice que tiene una distribución de Poisson de parámetro λ positivo y se escribe X$\stackrel{D}{=}$P(λ) si toma como valores todos los enteros no negativos y su función de cuantía viene dada por:

$$P(X = r) = e^{-?}\frac{?^{r}}{r!}$$

Observar que $\sum_{r=0}^{+\infty} P(X = r) = \sum_{r=0}^{+\infty} e^{-?}\frac{?^{r}}{r!} = e^{-?}\sum_{r=0}^{+\infty}\frac{?^{r}}{r!} = e^{-?}e^{?} = 1$

La variable de Poisson mide el número de sucesos que ocurren en un intervalo de longitud fija.

La distribución de Poisson puede también considerarse como el límite de una distribución B(n, $p = \frac{?}{n}$) cuando n crece indefinidamente. Comprobemos esta afirmación.

Sea X la variable aleatoria que mide el número de veces que ocurre un suceso en un intervalo de longitud L. Supongamos que el intervalo se divide en n subintervalos de manera que pueda decirse que es cero la probabilidad de que en uno de esos subintervalos ocurran dos sucesos. De este modo X puede considerarse como una binomial de parámetros n y $\frac{?}{n}$.

$$P(X = r) = \binom{n}{r}\left(\frac{?}{n}\right)^r \left(1-\frac{?}{n}\right)^{n-r} \Rightarrow \lim_{n\to+\infty} P(X = r) = \lim_{n\to+\infty} \binom{n}{r}\left(\frac{?}{n}\right)^r \left(1-\frac{?}{n}\right)^{n-r} =$$

$$= \lim_{n\to+\infty} \frac{n!}{r!(n-r)!}\left(\frac{?}{n}\right)^r \left(1-\frac{?}{n}\right)^{n-r} = \frac{?^r}{r!} \lim_{n\to+\infty} \frac{n(n-1).....(n-r+1)}{n^r\left(1-\frac{?}{n}\right)^r}\left(1-\frac{?}{n}\right)^n =$$

$$= \frac{?^r}{r!} \lim_{n\to+\infty} \frac{n}{n-?}\,\frac{n-1}{n-?}\,.......\frac{n-r+1}{n-?}\left(1-\frac{?}{n}\right)^n = \frac{?^r}{r!}e^{-?}$$

. Luego tenemos que $P(X = r) = \frac{l^r}{r!}e^{-\lambda}$ $r \in Z^+$ y la función de distribución es

$$F(x) = \sum_{r\le x} e^{-?}\frac{?^r}{r!} \quad \forall x \in R.$$

<u>Esperanza y Varianza de la distribución de Poisson.</u>

Usando la función generatriz de momentos se tiene:

$$?(t) = E\left[e^{tx}\right] = \sum_{r=0}^{+\infty} e^{tr}\frac{?^r}{r!}e^{-?} = e^{-?}\sum_{r=0}^{+\infty}\frac{(e^t?)^r}{r!} = e^{-?}e^{?e^t} = e^{?(e^t-1)}$$

$$?'(t) = ?e^t e^{?(e^t-1)} \Rightarrow ?'(0) = ? \Rightarrow E[X] = ?$$

$$?''(t) = ?e^t e^{?(e^t-1)} + ?^2 e^{2t} e^{?(e^t-1)} \Rightarrow ?''(0) = ? + ?^2 \Rightarrow Var[X] = ?$$

Para aproximar una B(n, p) por una P(λ) hay que tener en cuenta los siguientes aspectos.

$E[X_B] = np$, $Var\,[X_B] = npq$

$E[X_P] = \lambda = Var\,[X_P]$

De ellos se deduce que para que la aproximación tenga sentido λ=np y q tiene que ser un valor muy próximo a 1. En la práctica la aproximación es buena cuando p<0.1 y np≤5.

EJEMPLO I.

Las llamadas por avería en un puesto de servicio siguen una distribución de Poisson de media dos averías por semana. Calcular:

i) Probabilidad de que en una semana no haya ninguna avería.
ii) Probabilidad de que en una semana haya menos de cinco averías.
iii) Probabilidad de que haya seis averías en un mes.

<u>Solución.</u>

Se define la v.a X= “número de averías por semana” $\overset{D}{=}$ P(2).

i) $P(X = 0) = e^{-2} = 0.14$
ii) $P(X<5) = P(X \le 4) = 0.947$

iii) Considerando que un mes tiene cuatro semanas trabajaremos con la variable Y =" número de averías por mes" $\stackrel{D}{=}$P(8)
P(Y = 6)=0.122

EJEMPLO II.

La probabilidad de reacción negativa ante un fármaco de un individuo internado en un hospital es 0.05. Calcular la probabilidad de que en un total de 100 individuos

i) tres presenten reacción.
ii) No presente reacción ninguno.
iii) Presenten reacción mas de dos.

Solución.

La v.a. X = " número de individuos que presentan reacción "$\stackrel{D}{=}$B(100, 0.05)≈P(5)

i) P(X =3)=0.1404
ii) P(X = 0) = 0.0067
iii) $P(X>2) = 1- P(X \leq 2) = 1- 0.125 = 0.875$

EJEMPLO III.

Se sabe que una central telefónica atiende un promedio de 8 llamadas por minuto. Con los medios técnicos de que dispone la central, se pueden atender como máximo 12 llamadas por minuto, produciéndose una sobresaturación en la línea si se sobrepasa dicho número.

i) Hallar la probabilidad de que en un determinado minuto haya saturación.
ii) Usar la desigualdad de Tchebyshev para acotar la probabilidad de que el número de llamadas esté en el intervalo [μ-2σ, μ+2σ]

Solución.

i) Se define la variable X = " número de llamadas por minuto" $\stackrel{D}{=}$P(8)
$P(X>12) = 1-P(X\leq12)=1-0.9362=0.0638.$

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

$\mu = E[X] = 8 \quad \sigma = Des[X] = \sqrt{8} = 2.83 \Rightarrow P(2.34<X<13.66) \geq \frac{3}{4}$

Luego la probabilidad de que el número de llamadas recibidas por minuto esté entre dos y trece es mas del 75 %.

EJEMPLO IV.

La probabilidad de que una persona muera debido a un cierto virus es 0.001. ¿Cuál es la probabilidad de que mueran al menos tres en una población de 3000 personas afectadas?

Solución.

La v.a. X = " número de personas afectadas por el virus en una población de 3000" $\stackrel{D}{=}$B(3000, 0.001)≈P(3)

$P(X\geq3) = 1- P(X\leq2) = 1-0.4232=0.5768$

NOTA. Cuando una distribución de frecuencias cumple las condiciones que permiten suponer que sigue una ley de Poisson, la distribución teórica que mejor se ajusta es una Poisson que tiene la misma media que la distribución empírica.

.- Otras distribuciones discretas:

Geométrica .

En un proceso de Bernouilli se considera una variable X que cuenta el número de pruebas hasta la obtención del primer éxito. Si p es la probabilidad de éxito la función de cuantía de dicha variable sería:

$$P(X = k) = (1-p)^k p \quad k = 1, 2, 3, \ldots$$

$$E[X] = \frac{1}{p} \quad \mathrm{Var}\,[X] = \frac{1-p}{p^2}$$

Binomial Negativa.

En un proceso de Bernouilli se considera una variable X que cuenta el número de fracasos antes de que aparezca el n-ésimo éxito. La función de cuantía será en este caso:

$$P(X = k) = \binom{n+k+1}{k}(1-p)^k p^n \quad k = 0, 1, 2, 3, \ldots$$

$$E[X] = \frac{n(1-p)}{p} \quad \mathrm{Var}\,[X] = \frac{n(1-p)}{p^2}$$

Hipergeométrica.

Es el modelo mas adecuado para seleccionar una muestra de tamaño n, procediendo sin reposición, de un lote de tamaño N en el que existen Np elementos defectuosos. Por esto esta distribución se utiliza para elaborar planes de muestreo para aceptación en control de calidad.

La probabilidad de que una muestra de tamaño n contenga k elementos defectuosos viene dada por:

$$P(X = k) \frac{\binom{Np}{k}\binom{N-Np}{n-k}}{\binom{N}{n}} \quad k = 0, 1, 2, \ldots \;\; N = 1, 2, 3, \ldots \quad n = 1, 2, \ldots N.$$

$$E[X] = np \quad \mathrm{Var}\,[X] = npq$$

2.5.- Algunos ejemplos de distribuciones continuas.

- Distribución Exponencial. l

En un proceso de Poisson se considera una variable aleatoria continua T que representa el tiempo que transcurre entre la ocurrencia de dos sucesos consecutivos.

Sea μ el promedio de sucesos ocurridos en un intervalo temporal $(0, t_0)$ y sea X una variable de Poisson de parámetro μ. La probabilidad de que en el intervalo $(0, t_0)$ no haya ocurrido ningún suceso es $P(X = 0) = P(T > t_0) = e^{-\mu}$

Sea $? = \frac{\mu}{t_0}$ la media de sucesos por unidad de tiempo $\Rightarrow P(T > t_0) = e^{-\lambda t_0}$

$F(t_0) = P(T \leq t_0) = 1- P (T > t_0) =1- e^{-\lambda t_0}$

Luego T es una variable aleatoria cuya función de distribución viene dada como sigue.

$$F(t) = \begin{cases} 1-e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Se dice que T tiene una distribución exponencial de parámetro λ y escribiremos $T \overset{D}{=} \exp(\lambda)$.

Derivando la función de distribución se obtiene la función de densidad.

$$f(t) = \begin{cases} \lambda\ e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

<u>Esperanza y Varianza dela distribución exponencial .</u>

Como ya mencionamos anteriormente

$$E[T] = \int_{-\infty}^{+\infty} t\ f(t)\ dt = \int_{0}^{+\infty} t\ \lambda\ e^{-\lambda t} dt = \{\text{integrand o por partes}\} = \frac{1}{\lambda}$$

$$E[T^2] = \int_{-\infty}^{+\infty} t^2\ f(t)\ dt = \int_{0}^{+\infty} t^2\ \lambda\ e^{-\lambda t} dt = \{\text{integrand o por partes}\} = \frac{2}{\lambda^2}$$

$$Var[T] = \frac{1}{\lambda^2}$$

EJEMPLO I.

Sea T una v.a. que mide el tiempo de vida en horas de un cierto tipo de tubos de vacío. Supongamos que $T \overset{D}{=} \exp(0.001)$. El fabricante desea garantizar los tubos para una duración de t_0 horas. Determinar t_0 de modo que $P(T > t_0) =0.95$.

<u>Solución</u>.

$$P (T > t_0) = 1- F(t_0) = e^{-\lambda t_0} = 0.95 \Rightarrow -\lambda\ t_0 = \log 0.95 \Rightarrow t_0 = \frac{-\log 0.95}{\lambda} \Rightarrow$$

$t_0 = 51.25$ horas

EJEMPLO II.

Se ha comprobado que la duración de vida de ciertas componentes electrónicas sigue una distribución exponencial de media 8 meses.

i) Calcular la probabilidad de que la vida de una componente esté entre tres y doce meses.
ii) Calcular el percentil 0.95 de la distribución.
iii) Calcular la probabilidad de que una componente que ha vivido 10 meses viva 25 meses mas.

<u>Solución.</u>

Sea T la v.a. que mide la duración de vida en meses. $T \overset{D}{=} \exp(\frac{1}{8})$.

i) $P(3 \leq T \leq 12) = \int_3^{12} \frac{1}{8} e^{-\frac{1}{8}t} dt = - e^{-\frac{1}{8}t} \Big|_3^{12} = 0.47$

ii) $P(T \leq p) = 0.95$

$$P(T \leq p) = 1- e^{-\frac{p}{8}} = 0.95 \Rightarrow e^{-\frac{p}{8}} = 0.05 \Rightarrow - \frac{p}{8} = \log 0.05 \Rightarrow p = 23.97$$

iii) $P(T >25 / T >10) = \frac{P(\{T > 25\} \cap \{T > 10\})}{P(T > 10)} = \frac{P(T > 25)}{P(T > 10)} = \frac{e^{-25\frac{1}{8}}}{e^{-10\frac{1}{8}}} = e^{-15\frac{1}{8}}$

Observar que en general la distribución exponencial tiene la siguiente propiedad.

$P(T > t+s / T > t)=$

$$= \frac{P(\{T > t+s\} \cap \{T > t\})}{P(T > t)} = \frac{P(T > t+s)}{P(T > t)} = \frac{e^{-l(t+s)}}{e^{-lt}} = e^{-ls} = P(T > s)$$

Esta propiedad es la que sugiere su aplicación en ciertos problemas relativos a duración de vida.

Distribución Normal.

Se dice que una v.a. continua sigue una distribución normal de parámetros μ y σ y lo representaremos $X \overset{D}{=} N(\mu , \sigma)$ si su función de densidad es de la forma siguiente.

$$f(x) = \frac{1}{s\sqrt{2p}} e^{-\frac{(x-m)^2}{2s^2}} \qquad x \in (-\infty, +\infty)$$

$E[X] = \mu$ y $Var[X] = \sigma^2$.

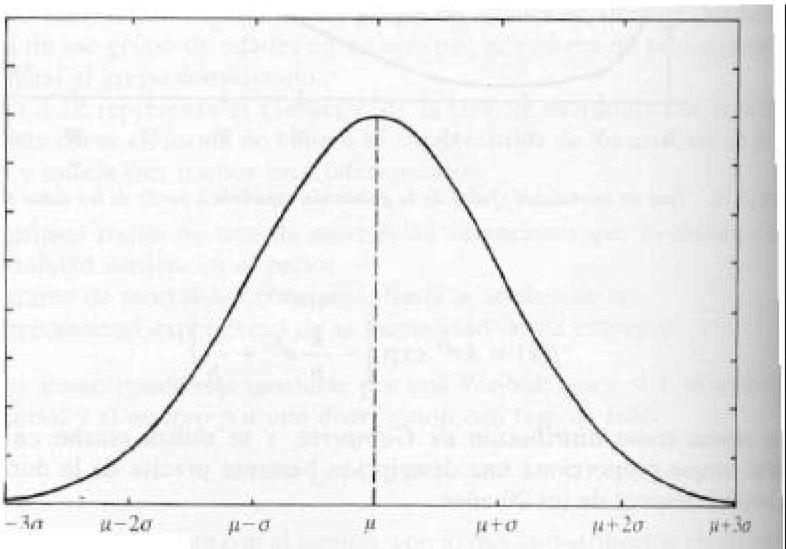

FIGURA 3.13. *Distribución normal.*

Observar que la variable $Z = \frac{X-\mu}{\sigma} \overset{D}{=} N(0,1)$.

Su función de densidad es $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ con $x \in (-\infty, +\infty)$.

Su función de distribución es $\varphi(z) = \int_{-\infty}^{z} f(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$ que es la única que tenemos tabulada.

Par calcular F(x) hay que tener en cuenta que el suceso $\{ X \le x \}$ es el mismo que el suceso $\{Z \le \frac{x-\mu}{\sigma}\}$, y por tanto $F(x) = P(X \le x) = P(Z \le \frac{x-\mu}{\sigma}) = \phi(\frac{x-\mu}{\sigma})$.

A continuación se muestra como se calcula la probabilidad de un intervalo.

i)

$$P(a \le X \le b) = P(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}) = \phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma}) .$$

ii)

$$P(|X| \le a) = P(-a \le X \le a) = P(\frac{-a-\mu}{\sigma} \le Z \le \frac{a-\mu}{\sigma}) = \phi(\frac{a-\mu}{\sigma}) - \phi(\frac{-a-\mu}{\sigma}) = 2\phi(\frac{a-\mu}{\sigma}) - 1$$

A partir de la distribución normal tipificada se puede calcular cualquier otra distribución normal.

EJEMPLO I.

Si $Z \overset{D}{=} N(0,1)$, determinar:

i) $P(Z \le 2.44)$
ii) $P(1 \le Z \le 3)$
iii) $P(Z > 1)$

Solución.

i) $P(Z \le 2.44) = 0.9927$

ii) $P(1 \leq Z \leq 3) = \phi(3)-\phi(1) = 0.9987-0.8413 = 0.1574$
iii) $P(Z > 1) = 1- P(Z \leq 1) = 1 - \phi(1) = 1- 0.8413 = 0.1587$

EJEMPLO II.

Sea $X \overset{D}{=} N(0.8,2)$. Calcular.

i) $P(X \leq 2.44)$
ii) $P(1 \leq X \leq 3)$
iii) $P(X > 1)$

Solución.

i) $P(X \leq 2.44) = P(Z \leq \frac{2.44 - 0.8}{2}) = P(Z \leq 0.82) = \phi(0.82) = 0.7939$

ii) $P(1 \leq X \leq 3) = P(\frac{1-0.8}{2} \leq Z \leq \frac{3-0.8}{2}) = f(1.1) - f(0.1) = 0.8643-0.5398$
$= 0.3245$

iii) $P(X > 1) = 1- P(X \leq 1) = 1- P(Z \leq \frac{1-0.8}{2}) = 1-\phi(0.1) =$
$= 1-0.5398 = 0.4602$

EJEMPLO III.

Calcular el valor de a tal que $P(Z \leq a) = 0.90$

Solución.

Si miramos las tablas de la normal tipificada tenemos que $P(Z \leq 1.28) = 0.8997$ y $P(Z \leq 1.29) = 0.90147$. Convendremos que entre esos dos valores tomaremos el que da un resultado más próximo a 0.90 que es $a=1.28$.

EJEMPLO IV.

Calcular el valor de a tal que $P(Z>a) = 0.85$

Solución.

$P(Z > a) = 1- P(Z \leq a) = 0.85 \Rightarrow P(Z \leq a) = 1 - 0.85 = 0.15$.
En las tablas observamos que $P(Z \leq -1.04) = 0.1492$ y $P(Z \leq -1.03) = 0.1515$.
Con el criterio del ejemplo anterior tomamos $a = -1.04$

EJEMPLO V.

Calcular a tal que $P(|Z| \leq a) = 0.99$

Solución.

$P(|Z| \leq a) = P(-a \leq Z \leq a) = 2\phi(a) - 1 = 0.99 \Rightarrow \phi(a) = \frac{1.99}{2} = 0.9950 \Rightarrow a = 2.58$

EJEMPLO VI.

Probar que si $X \overset{D}{=} N(\mu, \sigma)$ se verifica:

i) $P(\mu- \sigma < X < \mu+ \sigma) = 0.6826$
ii) $P(\mu- 2\sigma < X < \mu+ 2\sigma) = 0.9546$
iii) $P(\mu- 3\sigma < X < \mu+ 3\sigma) = 0.9974$

Solución.

i) $P(\mu - \sigma < X < \mu + \sigma) = P\left(\frac{\mu - s - \mu}{s} < Z < \frac{\mu + s - \mu}{s}\right) = P(-1 < Z < 1) =$

$= 2\phi(1) - 1 = 2 \times 0.8413 - 1 = 0.6826$

ii) $P(\mu - 2\sigma < X < \mu + 2\sigma) = P\left(\frac{\mu - 2s - \mu}{s} < Z < \frac{\mu + 2s - \mu}{s}\right) = P(-2 < Z < 2) =$

$= 2\phi(2) - 1 = 2 \times 0.9773 - 1 = 0.9546$

iii) $P(\mu - 3\sigma < X < \mu + 3\sigma) = P\left(\frac{\mu - 3s - \mu}{s} < Z < \frac{\mu + 3s - \mu}{s}\right) = P(-3 < Z < 3) =$

$= 2\phi(3) - 1 = 2 \times 0.9987 - 1 = 0.9974$

De los resultados del ejemplo anterior se desprende que al observar un gran número de valores de una v.a. normal cabe esperar que alrededor de las 2/3 partes de los datos estén en el intervalo $(\mu - \sigma, \mu + \sigma)$, alrededor de 95% en $(\mu - 2\sigma, \mu + 2\sigma)$ y el 99% entre $(\mu - 3\sigma, \mu + 3\sigma)$. Esto indica que si un conjunto grande de datos no responde a esta situación es porque no siguen una distribución normal.

EJEMPLO VII.

Una de las primeras aplicaciones de la curva normal la hizo Bessel que en 1818 comprobó que los errores de medida de 300 observaciones astronómicas coincidían con bastante aproximación con los previstos por Gauss en la curva normal. Suponiendo que la v.a. que mide estos errores tiene media 0 y una desviación típica de 4 grados, calcular:

i) La probabilidad de que el error en una medida no sea mayor de seis grados.
ii) La probabilidad de que se cometa un error por defecto mayor que ocho grados.
iii) Si llamamos grandes a los errores mayores que siete grados, calcular el numero esperado de errores grandes en 300 observaciones.

Solución.

i) $P(|X| \leq 6) = P(-6 \leq X \leq 6) = P\left(-\frac{6}{4} \leq Z \leq \frac{6}{4}\right) = 2f(1.5) - 1 = 0.8663$

ii) $P(X \leq -8) = P\left(Z \leq -\frac{8}{4}\right) = f(-2) = 0.02275$

iii) $P(|X| > 7) = 1 - P(|X| \leq 7) = 1 - 0.91988 = 0.08012$

Si se considera cada observación como una prueba de Bernouilli en la que el éxito es obtener un error grande, la variable aleatoria que mide el número de errores grandes en 300 observaciones es una $B(300, 0.08012)$ y su esperanza es $np = 300 \times 0.08014 = 24$ que es número esperado de errores grandes en 300 observaciones.

2.6.- La Normal como aproximación de otros distribuciones.

Teorema Central del Límite.

Sean X_1, $X_{2,\ldots,}$ X_n variables aleatorias independientes con medias $\mu_1, \mu_2, \ldots, \mu_n$ y desviaciones $\sigma_1, \sigma_2, \ldots, \sigma_n$ respectivamente y con una distribución cualquiera no necesariamente la misma. Sea $Y = X_1 + X_2 + \ldots + X_n$ y sea $Z = \dfrac{Y - \sum_{i=1}^{n} \mu_i}{\sqrt{s_1^2 + s_2^2 + \ldots + s_n^2}}$.

Se demuestra que cuando n crece la variable Z tiende hacia una distribución N(0,1).

Este resultado implica que cuando n es grande se puede considerar que Y tiene una distribución $N(\sum_{i=1}^{n} \mu_i, \sqrt{\sum_{i=1}^{n} s_i^2})$.

Si todas las variables X_i tienen la misma media μ y la misma desviación σ en tonces Y se distribuye cuando n crece como una normal de media $n\mu$ y de desviación $\sqrt{n}\,s$

Observación.

El teorema Central del Límite es uno de los resultados mas importantes que se conocen en estadística. Su demostración requiere un nivel que supera con mucho al del curso, pero su aplicación es de gran interés. En el se basa la justificación teórica de la aproximación por la distribución normal de otras distribuciones tales como la Binomial y la Poisson .

<u>Aproximación de una distribución B(n, p) por una distribución Normal.</u>

Una variable $X \overset{D}{=} B(n, p)$ es la suma de n variables aleatorias X_i de Bernouilli independientes. $E[X_i] = p$ y $Var[X_i] = pq$ $\forall i = 1, 2, \ldots, n$. Aplicando el teorema central del límite se concluye que cuando n es grande la variable binomial de parámetros n y p puede considerarse como un normal de parámetros np y $\sqrt{npq}$ y por tanto la variable

$$Z = \frac{X - np}{\sqrt{npq}} \overset{D}{=} N(0,1)$$

En la práctica la aproximación es buena cuando n grande (mayor que 30) y p no está muy próximo a cero ni a uno, también cuando n es pequeño si p está próximo a 0.5. Es correcto utilizar la normal como aproximación de la binomial cuando $np \geq 5$ y $nq \geq 5$. Si p está próximo a 0.5 la aproximación es válida si $np > 3$ aún cuando n sea pequeño.

Cuando se realiza la aproximación hay que tener en cuenta que se está aproximando una distribución discreta por una continua y eso obliga a hacer ciertas correcciones que se presentan a continuación.

$P (X_B = a) = P (a - 0.5 \leq X_N \leq a + 0.5)$

$P (a \leq X_B \leq b) = P (a - 0.5 \leq X_N \leq b + 0.5)$

EJEMPLO I.

Una variable aleatoria X sigue una distribución B(20,0.45). Calcular P(X=10) y $P(3 \leq X \leq 6)$:

i) Usando las tablas de la binomial.
ii) Usando una N(9, 2.22)

<u>Solución</u>

i) $P(X=10) = 0.1593$

$P(3\leq X\leq 6) = F(6)- F(3) + P(X=3) =0.1299 - 0.0049 + 0.0040 = 0.129$

ii) $P (9.5<X<10.5) = P(0.23<Z<0.68) = \phi(0.68)-\phi(0.23) =$
$0.7517-0.5910= 0.1607$

$P(2.5\leq X\leq 6.5) =P(-2.93 \leq Z \leq -1.13) = P(1.13\leq Z \leq 2.93) =$
$\phi(2.93)-\phi(1.13) = 0.998305-0.8708=0.127505$

La tabla siguiente presenta los resultados que muestran la bondad de la aproximación.

	B(20, 0.45)	N(9,2.22)
P(X=10)	0.1593	0.1607
P(3≤X≤6)	0.129	0.127505

EJEMPLO II.

Un profesor decide hacer un examen en forma de test con un cuestionario de 100 preguntas. Cada pregunta va acompañada de cinco respuestas de las que sólo una es la correcta. Calcular la probabilidad de que un alumno , que responde al azar una entre las cinco respuestas, obtenga entre 10 y 30 respuestas correctas.

Solución.

La variable aleatoria que mide el número de respuestas acertadas en el cuestionario es una B(100, 1/5).Como np=100x0.2=20 y nq=100x0.8=80 se puede hacer la aproximación por una N(20,4).

$P(10 \leq X_B \leq 30) \cong P(9.5 \leq X_N \leq 30.5) = P(-2.63 \leq Z \leq 2.63) =$
$= 2\phi(2.63)-1=2x0.995731 – 1 = 0.991462$

Aproximación de una distribución P(l) por una distribución Normal.

Una distribución de Poisson de parámetro λ>5 se puede aproximar por una distribución normal de media λ y de desviación $\sqrt{?}$

EJEMPLO I.

En un proceso de fabricación de película fotográfica aparece por término medio 1 defecto cada 20 metros. Si la distribución de defectos es una Poisson, calcular la probabilidad de 6 defectos en un rollo de 200 m

i) Directamente.
ii) Aproximando por una normal.

Solución.

i) Se pide P(X = 6) siendo $X \overset{D}{=} P(10)$.

$$P(X = 6) = e^{-10}\frac{10^6}{6!} = \mathbf{0.0631}$$

ii) $X \overset{D}{=} P(10)\approx N(10, 3.16)$

$P(X_P = 6) \cong P(5.5 \leq X_N \leq 6.5) = P(-1.42 \leq Z \leq -1.11) = \phi(1.42)-\phi(1.11)=$
$= 0.9222 – 0.8665 = \mathbf{0.0557}$

2.7.- Distribuciones asociadas a la distribución Normal.

- c^2 de Pearson.

Sean $X_1, X_2, \ldots, X_n$ variables aleatorias todas ellas con distribución N(0, 1) e independientes. Se define la variable Ji-cuadrado de Pearson con n grados de libertad y se representa $?^2_n$, como:

$$?^2_n = X_1^2 + X_2^2 + \ldots + X_n^2$$

$E[X_i^2] = 1$ ya que $E[X_i] = 0$ y $Var[X_i] = E[X_i^2] - (E[X_i])^2$ entonces

$E[?^2_n] = E[X_1^2 + X_2^2 + \ldots + X_n^2] = E\left[X_1^2\right] + E\left[X_{21}^2\right] + \ldots + E\left[X_n^2\right] = n$

Por un procedimiento mas complejo se demuestra que

$Var[?^2_n] = 2n$

La función de distribución se obtiene a partir de la tabla que da colas a la derecha de un punto es decir $P(?^2 \geq ?^2_{a,n}) = a$.

EJEMPLO I.

El valor crítico $?^2_{a,n}$ en los siguientes casos es:

i) $\alpha = 0.05$ y $n = 6 \Rightarrow ?^2_{0.05,6} = 12.592$

ii) $\alpha = 0.90$ y $n = 6 \Rightarrow ?^2_{0.90,6} = 2.204$

- F de Fisher.

Se llama v.a. de Fisher con (n, m) grados de libertad a $F_{n,m} = \dfrac{?^2_n}{?^2_m}$

Las tablas indican que $P(F_{n,m} > F_{\alpha,n,m}) = \alpha$, es decir proporcionan los valores $F_{\alpha,n,m}$ que dejan a su derecha un área α (0.05 y 0.01).

EJEMPLO 1.

Comprobar usando las tablas correspondientes que $F_{\alpha,n,m}$ siendo α=0.05, n=3 y m=15 es 3.2874 y si α=0.01, n=5 y m=2 es 99.299

- t de Student.

Sea Z = N(0, 1) y $?^2_n$. Se llama v.a. t de Studen con n grados de libertad a la variable que se define como sigue:

$$t_n = \frac{Z}{\left[\frac{1}{n} ?^2_n\right]^{\frac{1}{2}}}$$

La función de densidad es una curva simétrica respecto de OY. Para valores grandes de n la curva se aproxima mucho a la curva normal tipificada.
La tabla proporciona $P(t_n \geq t_{\alpha,n}) = \alpha$.

EJEMPLO I.
El valor $t_{\alpha,n}$ cuando $n = 5$ y $\alpha = 0.10$ es 1.476