Tema 3 **VARIABLE ALEATORIA** y **DISTRIBUCIONES DE PROBABILIDAD**

3.1. INTRODUCCIÓN

3.2. TEORÍA COMBINATORIA

3.3. CÁLCULO DE PROBABILIDADES

3.4. VARIABLE ALEATORIA

3.5. DISTRIBUCIONES DE PROBABILIDAD

Distribuciones de probabilidad discretas

Distribuciones de probabilidad continuas

Aproximación entre las distribuciones binomial, Poisson y normal

VARIABLE ALEATORIA y DISTRIBUCIONES DE PROBABILIDAD

Resumen: La inmensa mayoría de los procesos de la salud son fenómenos aleatorios, es decir, de resultados inciertos.

El **Cálculo de probabilidades** permite conocer el grado de seguridad que tenemos acerca de determinados resultados de tales procesos.

El concepto de **variable aleatoria** permite conocer el tipo de fenómeno, probabilidad de ciertos resultados y parámetros de interés como son la media, la varianza,…, del conjunto de resultados numéricos.

Este conocimiento es posible tanto si la probabilidad de los distintos resultados se conoce de forma empírica, como si la regularidad de tales fenómenos permite ajustarle una **distribución de probabilidad** bien discreta o bien continua.

3.1. INTRODUCCIÓN

El fin principal de la Estadística es utilizar los datos de que se dispone para extraer conclusiones sobre una variable observada en una población a la que normalmente no podremos acceder de manera completa.

En general, dispondremos de información parcial que obtenemos eligiendo al azar algunos individuos de la población para formar la muestra, lo que constituye un experimento aleatorio.

La Teoría combinatoria permite calcular cuántas y qué muestras son posibles y constituye una herramienta básica del *Cálculo de probabilidades*, que se encarga de estudiar ese tipo de experimentos y de proporcionar modelos para las *distribuciones de probabilidad* de algunas variables importantes.

Pretendemos en este tema hacer un breve recordatorio de estas materias que nos van a permitir: codificar todo tipo de artículos, formar equipos de trabajo, "combinar" productos para ofertas, etc.

El Cálculo de probabilidades nos permitirá medir el grado de seguridad que tendremos acerca de que se produzcan determinados resultados tras el lanzamiento de campañas publicitarias, tras la realización de muestreos, etc.

El concepto de **variable aleatoria** y sus **distribuciones de probabilidad** van a facilitar el conocimiento de procesos aleatorios, así como el cálculo de parámetros de interés, (media, varianza, ...), y probabilidad de determinados resultados.

3.2. TEORÍA COMBINATORIA

Es una rama de las Matemáticas que permite resolver dos problemas:

1º. Cuántos grupos de "n" objetos pueden formarse a partir de "m" objetos dados
2º. Cómo formar dichos grupos

Se consideran tres tipos diferentes de grupos:

- PERMUTACIONES, todos los grupos están formados por todos los objetos de partida. Dos grupos se diferencian, solamente, en el orden de colocación de los objetos.

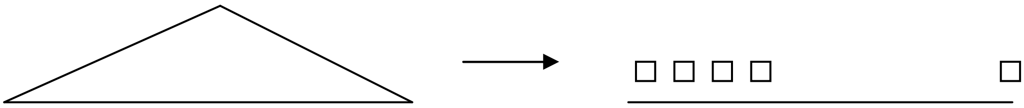
- COMBINACIONES, los grupos están formados por solo n objetos de los m dados y dos grupos se diferencian en la composición, no importando el orden de colocación.

- VARIACIONES, también los grupos están formados por n de los m objetos dados. Dos grupos se diferencian porque tienen algún elemento diferente o bien cuando, teniendo los mismos, están colocados en diferente orden.

En los tres tipos de grupos, los elementos pueden o no repetirse, clasificándose entonces los grupos en permutaciones, combinaciones y variaciones con repetición o simples. La designación del tipo de grupo y la fórmula para obtener el número de grupos se recoge en la siguiente tabla:

	Simples	Con repetición
PERMUTACIONES	$P_n = n!$	$PR_n^{a,b,c,...} = \dfrac{n!}{a!b!c!...}$

¿Cómo diferenciar los distintos grupos?



m objetos n objetos

1.- ¿Pueden repetirse los objetos?

sí $\rightarrow$ con repetición

no $\rightarrow$ sin repetición

hecho que sólo afecta a las fórmulas a emplear

2.- ¿Intervienen <u>necesariamente</u> todos los objetos en todos los grupos?

sí $\rightarrow$ **PERMUTACIONES**

no $\rightarrow$ es preciso contestar a una tercera pregunta

3.- ¿Influye el orden de colocación?

sí $\rightarrow$ **VARIACIONES**

no $\rightarrow$ **COMBINACIONES**

nota final:

$\binom{m}{n}$ se llama **número combinatorio** "m sobre n".

Tiene aplicaciones interesantes como son los coeficientes en desarrollo de potencias, partes de un conjunto, número de coaliciones que un jugador puede formar, etc. Entre sus propiedades más interesantes están:

1ª.- $\binom{m}{0} = \binom{m}{m} = 1$ Propiedades que permiten construir el triángulo de Tartaglia:

2ª.- $\binom{m}{n} = \binom{m}{m-n}$

1 1

3ª.- $\binom{m}{n} + \binom{m}{n+1} = \binom{m+1}{n+1}$

1 2 1

3.3. CÁLCULO DE PROBABILIDADES

Rama de las Matemáticas que tiene por finalidad dar una medida del grado de seguridad que se tiene acerca de los diferentes resultados de un experimento aleatorio.

A cada uno de tales resultados lo llamaremos **suceso elemental** y al conjunto de todos lo sucesos elementales **espacio muestral** (E).

Un **suceso** (A) es cualquier subconjunto del espacio muestral y son sucesos extremos el **suceso imposible** (Φ), que nunca puede ocurrir y el **suceso seguro** (E), que siempre ocurre.

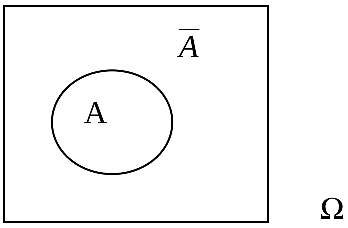
El espacio de sucesos (Ω) es el conjunto formado por todos los sucesos posibles de un experimento aleatorio coincide con el conjunto de las partes de E, que designaremos P(E) y su cardinal es :

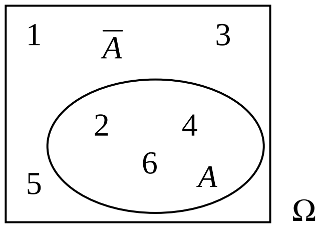
$$\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n} = 2^n$$

Si por ejemplo el experimento aleatorio consiste en lanzar un dado:

Un **suceso elemental** es, por ejemplo, sacar un tres.
El **espacio muestral** es E = { 1, 2, 3, 4, 5, 6 }
Un **suceso** A, sacar un número par.
Un **suceso imposible**, sacar un siete y un **suceso seguro** sacar un número menor que siete.
El **espacio de sucesos** es $\Omega = \{\Phi, 1, 2, ..., 6, 12, 13, ..., 56, 123, ..., 123456\}$ que tiene $2^6 = 64$ sucesos,(aquí Φ representa al conjunto vacío). Sobre Ω se pueden

Gráficamente mediante diagramas de Venn:



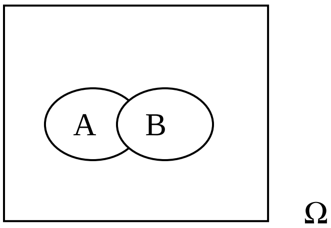


Ejemplo: A.- sacar un número par

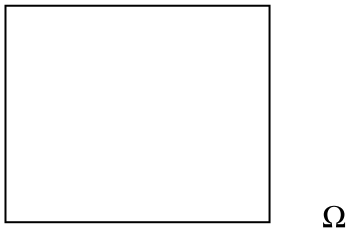
$\overline{A}$.- sacar un número impar

La **unión** de dos sucesos A y B es un suceso que tiene lugar cuando acontece al menos uno de ellos. Lo designaremos A$\cup$B.

Gráficamente:



Ejemplo:

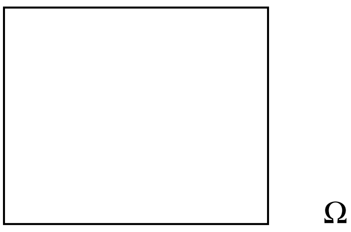


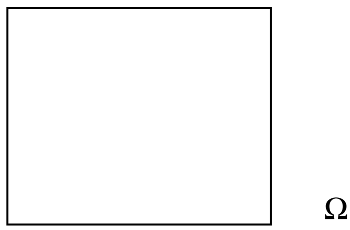
Ejemplo[1]: A.- número par

B.- número primo

La **intersección** (A $\cap$ B) es otro suceso que acontece cuando suceden simultáneamente A y B.

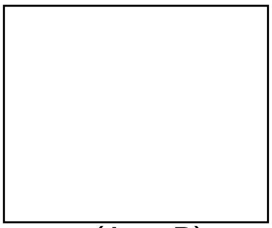
Gráficamente:





La **diferencia** (A-B) entre A y B es un suceso que acontece cuando sucede A pero no B y la **diferencia simétrica** (A Δ B) cuando ocurre uno de los sucesos A o B pero no los dos simultáneamente.

A – B



(A Δ B)

En la práctica estaremos interesados en calcular la probabilidad de algunos sucesos de Ω para lo que se han propuesto diferentes modelos.

Definición clásica de probabilidad

La **ley empírica del azar** o ley de estabilidad de las frecuencias establece que, realizado un experimento aleatorio muchas veces, las frecuencias relativas de cada suceso tienden a un número fijo, que nos interesa definir como probabilidad de este suceso. La definición clásica de probabilidad, basada en este hecho es la siguiente:

La probabilidad de que aparezca determinado suceso es el cociente entre el número de casos favorables a ese suceso y el número total de casos siendo estos mutuamente simétricos.

Definición que se aplica en la práctica a la resolución de gran número de problemas mediante la fórmula:

p(A)= nº de casos favorables /nº de casos posibles), conocida como **regla de Laplace.**

Esta definición, tiene dos inconvenientes: exige equiprobabilidad y solo es aplicable

De Finetti que define la probabilidad **subjetiva**, donde la probabilidad mide el "grado de certeza" que para un cierto individuo tiene la ocurrencia de un suceso.

Definición axiomática de probabilidad

Establecida en 1930 por el matemático ruso **Kolmogorov**[2] se basa en la Teoría de la medida y proporciona una base sólida al concepto de probabilidad y al Cálculo de probabilidades.

De un modo general, puede decirse que la probabilidad es una función o una regla p que permite asignar a todo suceso A de un espacio de sucesos Ω un único número real llamado probabilidad de A que designamos p(A), que satisface los siguientes axiomas:

A1.- Para todo suceso A, es $p(A) \geq 0$

A2.- Si A y B son incompatibles, entonces $p(A \cup B) = p(A) + p(B)$

A3.- Si E es el espacio muestral, $p(E) = 1$

De estos axiomas se deducen algunas propiedades interesantes:

1.- La probabilidad de un suceso es un número comprendido entre 0 y 1

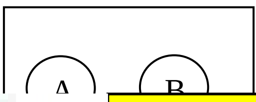
2.- La probabilidad del suceso imposible es nula: $p(\Phi) = 0$

3.- La probabilidad de un suceso es igual a uno menos la probabilidad del suceso contrario: $p(A) = 1 - p(\overline{A})$

4.- Si un suceso A está incluido en un suceso B ($A \subset B$), se cumple $p(A) \leq p(B)$.

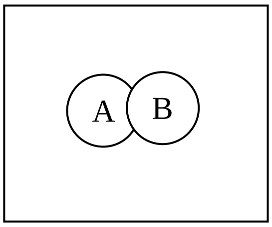
Probabilidad de la unión de dos sucesos, ($\cup$, ó, +)

a) Incompatibles, que no pueden ocurrir a la vez.



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

b) Compatibles, que pueden ocurrir a la vez



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

Generalización a tres o más sucesos

a) Incompatibles

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

b) Compatibles

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

Probabilidad de la interacción de dos sucesos, ($\cap$, y, •)

a) Independientes, el resultado del primero no influye en el segundo.

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

b) Dependientes, el resultado del primero influye en el segundo.

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

Probabilidad condicionada, de un suceso A, sabiendo que ha ocurrido el suceso B.

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

Generalización a tres o más sucesos

a) Independientes

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

Teorema de la probabilidad total

	B	$\overline{B}$
A_1		
A_2		
$\vdots$		
A_k		

Si los sucesos A_1, A_2, ..., A_k son incompatibles dos a dos y su unión es E, y por otra parte consideramos los sucesos incompatibles B y $\overline{B}$ cuya unión también es E, se verifica:

$$p(B) = p(A_1 \cap B) + p(A_2 \cap B) + \ldots + p(A_k \cap B)$$

Y si los sucesos A_i y B son dependientes:

$$p(B) = p(A_1)\, p(B/ A_1) + \ldots + p(A_k)\, p(B/ A_k)$$

Teorema de Bayes

Con los mismos supuestos anteriores:

$$p(A_i/B) = \frac{p(A_i \cap B)}{p(B)} = \frac{p(A_i)\, p(B/A_i)}{\sum_{i=1}^{k} p(A_i)\, p(B/A_i)}$$

Este teorema fue propuesto por Bayes[3] y ha sido el origen de la inferencia bayesiana. su aplicabilidad es muy grande en problemas de decisión, ya que permite

3.4. VARIABLE ALEATORIA

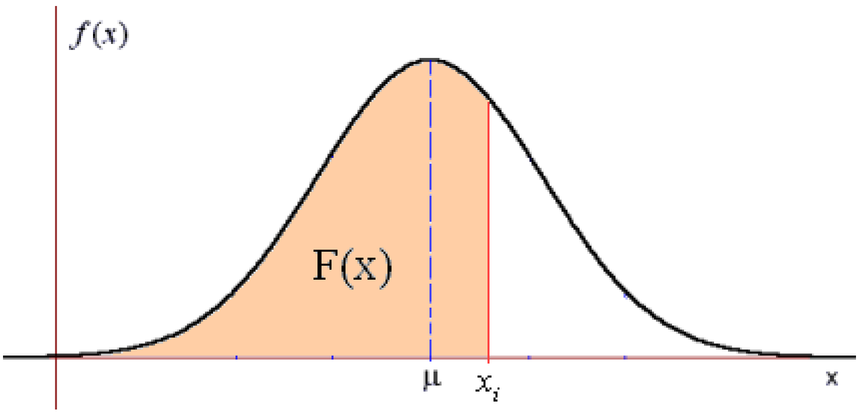
Una **variable aleatoria** o estocástica X (v.a. X) es una función que asocia a cada resultado de un experimento aleatorio x_i un número real que es su probabilidad:

$p(X = x_i) = p(x_i) = p_i$ las probabilidades p_i han de ser mayores o iguales a cero, y además, debe cumplirse que su suna sea la unidad, $\sum_i p_i = 1$

X queda absolutamente definida si conocemos **p(x)** que es la **función de cuantía** si X es discreta o **f(x), función densidad de probabilidad** si X es continua. A partir de p(x) o f(x) se define la correspondiente **función de distribución F(x)** de la forma:

$F(x)=p(X \leq x)$, probabilidad de que la variable tome valores menores o iguales que x

Gráficamente para una variable aleatoria X continua:



Los parámetros más interesantes son en ambos casos:

La media o esperanza matemática, la varianza y la desviación típica, que son parámetros de mucho interés en la toma de decisiones se obtienen mediante los siguientes cálculos.

Si la variable aleatoria es discreta:

Media.- $\mu = E[X] = \sum_{i=1}^{n} x_i p_i$, *suma de los productos de los valores por sus probabilidades*

Varianza.- $\sigma^2 = V(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p_i = \sum_{i=1}^{n} x_i^2 p_i - \mu^2 = E\left[x^2\right] - E[x]^2$,*cuadrados de los valores por sus probabilidades menos el cuadrado de la media*

Desviación típica.- $\sigma = +\sqrt{\sigma^2}$

Y para una v. a. continua:

Media.- $\mu = E[X] = \int_{x\,min}^{x\,max} x\, f(x)dx$

Varianza.- $\sigma^2 = V(X) = \int_{x\min}^{x\max} (x-\mu)^2 f(x)d(x) = \int_{x\min}^{x\max} x^2 f(x)d(x) - \mu^2$

Desviación típica.- $\sigma = +\sqrt{\sigma^2}$

La inmensa mayoría de los procesos o fenómenos económicos, biológicos, etc, son de resultados inciertos, el conocimiento del concepto de variable aleatoria así como el cálculo de parámetros y probabilidades nos ayuda a comprender mejor dichos fenómenos.

En unos casos las probabilidades de los distintos resultados podremos aproximarlas como consecuencia de experiencias pasadas, aunque la mayor parte de los fenómenos de la Naturaleza siguen exacta o aproximadamente unas pocas leyes bien conocidas que son llamadas **leyes o distribuciones de probabilidad teóricas**. Cada una de ellas es en realidad una familia de leyes que, teniendo la misma "forma", difieren unas de otras solo en sus parámetros (media y desviación típica generalmente).

En el siguiente apartado estudiamos con más detalle, aunque brevemente, las más

3.5. DISTRIBUCIONES DE PROBABILIDAD

Las distribuciones que siguen permiten analizar el comportamiento de gran cantidad de variables aleatorias que son modelos para explicar fenómenos reales en términos de poder calcular parámetros como la media, la varianza y desviación típica, así como las probabilidades de los resultados que interesen.

Distribuciones de probabilidad discretas

Binomial B(n, p)

Supongamos un **experimento dicotómico**, es decir, susceptible sólo de dos posibles resultados, llamados también de Bernouilli A y $\overline{A}$ (éxito / fracaso, cara / cruz, niño / niña…). Si designamos por **p** a la probabilidad de A y por q = 1-p la de $\overline{A}$, si se repite el experimento **n** veces sin que varíe a lo largo de todo el proceso la probabilidad p, la probabilidad de que el resultado A se presente r veces (r= 0, 1, 2,...,n) es la función de cuantía de la v. a. X, llamada binomial de parámetros n y p, y tiene por valor:

$$P(X = r) = \binom{n}{r} p^r q^{n-r}$$

Sus parámetros media, varianza y desviación típica son:

$$\mu = np, \quad \sigma^2 = npq \quad y \quad \sigma^2 = +\sqrt{npq}$$

Para n $\leq$10 las probabilidades están tabuladas (tabla 5).

Se utiliza en el control de calidad por atributos, en muestreo para aceptación de lotes de materias primas y en las técnicas de muestreo con reemplazamiento.

Poisson P(λ)

Es una distribución de probabilidad discreta que se ajusta muy bien a los siguientes experimentos aleatorios:

2. Repartos en medios continuos; clientes por zonas de una ciudad, bombas V2 sobre Londres durante la Segunda Guerra Mundial (GMII[4]), etc
3. Sucesos independientes en el tiempo; llegada de clientes a establecimientos, coches a estaciones de servicio, llamadas a centralitas, urgencias a hospitales, etc.

La v.a. X número de ocurrencias por unidad de tiempo de cualquiera de esta naturaleza sigue aproximadamente una distribución propuesta por Poisson de un único parámetro λ. La probabilidad de que se produzcan r ocurrencias, (r= 0, 1, 2, …, n), tiene por valor:

$$P(X=r) = e^{-\lambda}\frac{\lambda^r}{r!}$$, e = 2.7182...., base de los logaritmos neperianos

Sus parámetros media, varianza y desviación típica son:

$$\mu = \lambda, \quad \sigma^2 = \lambda \quad \text{y} \quad \sigma = +\sqrt{\lambda}$$

Probabilidades para valores λ iguales o menores de 10 están tabuladas (tabla 6).

Se utiliza en la simulación de procesos de llegadas en muchos fenómenos de espera y aparece con gran frecuencia en los estudios de fiabilidad.

Otras distribuciones discretas son:

La distribución **geométrica Ge(p)**, que analiza la probabilidad de que se produzcan r-1 fallos antes del primer éxito en un experimento de Bernouilli. Esta distribución fue propuesta por Pascal y se utiliza en estudios de fiabilidad.

La **binomial negativa BN(k, p)** en los mismos procesos analiza la probabilidad de que se produzcan r fallos antes del k-ésimo éxito.

La distribución **hipergeométrica HG(N, D, n), que** se utiliza en el muestreo de poblaciones finitas. N es el tamaño poblacional y la población consta de dos grupos de individuos de tamaños D y N-D respectivamente, se extraen para formar la muestra n

Distribuciones de probabilidad continuas

Normal N (μ,σ)

Es, con mucho, la más importante de las distribuciones de probabilidad continuas y ocupa un lugar central en el estudio de muchas ramas de investigación.

Su primera formulación matemática se debe a De Moivre en 1733. Muy vinculados a su historia se hallan otros matemáticos, particularmente Gauss[5], que la utilizó en 1808 en un famoso trabajo sobre el movimiento de cuerpos celestes.

Aparece en multitud de procesos y medidas en los campos de la Medicina, Biología, Física, Economía, etc. La razón de tal abundancia de variables normales la justifica el **teorema central del límite** una de cuyas lecturas es*: si una v. a. X* (pensemos por ejemplo en nuestra estatura*) toma valores como consecuencia de una gran cantidad de causas independientes* (…) *que actúan sumando sus efectos, siendo cada efecto individual de escasa importancia frente al resto, es de esperar que tales valores sigan una distribución normal*, y si los efectos fueran multiplicativos una log-normal o ley de Galton.

Depende de dos parámetros μ y σ que son la media y la desviación típica.

Gráficamente se la conoce con el nombre de campana de Gauss.

Están entre sus propiedades unas importantes:

- Es simétrica respecto a la recta x = μ
- En su centro coinciden la media, la mediana y la moda.
- Es unimodal con máximo el punto (μ, 1/($\sigma\sqrt{2\pi}$))
- Tiende asintomáticamente a cero cuando x→$\pm\infty$
- Tiene dos puntos de inflexión en x = $\mu \pm \sigma$
- El intervalo (μ-σ,μ+σ) contiene aproximadamente (~) el 68% de los datos.
- El intervalo (μ-1.96σ, μ+1.96σ) contiene ~ 95% de los datos
- El intervalo (μ-2σ, μ+2σ) contiene ~ 95.4% de los datos
- El intervalos (μ-3σ, μ+3σ) contiene ~ 99.7% de los datos
- Es reproductiva, la suma de k variables aleatorias normales es una normal de media suma de las medias y varianza suma de las varianzas.

La tipificación de los valores x de una variable normal X~ N (μ,σ) se obtiene:

$z = \dfrac{x-\mu}{\sigma}$, el valor z es el de una variable normal N (0, 1) llamada normal típica o tipificada cuya función de distribución está tabulada (tabla 1)

Otras distribuciones continuas son:

La distribución **uniforme U(a,b)** cuya gráfica es un rectángulo de base el intervalo (a,b) y área 1. La uniforme U(0,1) es una distribución básica en la generación de números aleatorios para simulación.

La distribución **exponencial E(λ)** es la distribución del tiempo que transcurre entre dos ocurrencias consecutivas de un fenómeno regido por la ley de Poisson.

Su función densidad de probabilidad es:

La distribución **gamma (a, p)** se ajusta al tiempo transcurrido entre las ocurrencias k y k+p de un proceso poissionano de media a (ley de Erlang).

Es la suma de p variables exponenciales y se utiliza en los mismos casos.

La distribución **beta (p, q)** puede ajustarse a muchas distribuciones empíricas mediante una adecuada elección de los parámetros p y q. Se utiliza en gestión de proyectos con la metodología PERT para la duración de actividades.

La distribución de **Weibull W (b, λ)** coincide con la exponencial cuando b = 1 y se utiliza para modelar el comportamiento ante la fatiga de diferentes materiales y como modelo de distribución en la teoría de la fiabilidad.

Distribuciones asociadas a la normal

Si bien no se ajustan a fenómenos reales, las tres distribuciones que siguen son básicas en Inferencia Estadística:

La distribución **chi-cuadrado** χ^2 de Pearson, de n grados de libertad es la distribución de una v.a. suma de los cuadrados de n v.a,s. normales N(0, 1) e independientes, su media es n y su varianza 2n. Es de enorme importancia en Inferencia Estadística; el método de mínimos cuadrados lleva frecuentemente a una distribución χ^2, interviene en la definición de otras distribuciones y es la base de contrastes de hipótesis de mucha utilidad. Su gráfica es asimétrica positiva y está tabulada en función de n y α nivel de significación (tabla 3)

La distribución **t de Student,** (seudónimo del químico Gosset[6] que la dedujo en 1908). Es una distribución función de otras dos X e Y ambas independientes entre sí; X es una N(0, 1) e Y una χ^2 con n grados de libertad.

Su gráfica tiene forma de campana simétrica platicúrtica y tiende a la normal al crecer n. Está tabulada en función de n y α (tabla 2)

La distribución **F de Snedecor** (F en honor a su maestro Fisher). Es la distribución de una v.a. cociente de dos distribuciones chi-cuadrado divididas cada una de ella por sus correspondientes grados de libertad n y m.

Juega un papel básico en el Análisis de la Varianza (ANOVA) y en el diseño de experimentos.

Está tabulada en función de n, m, y α (tabla 4 para distintos valores de α)

nota: en Inferencia Estadística se define:

(**grados de libertad**) = (tamaño muestral) - (parámetros estimados) ó
(tamaño muestral) - (relaciones impuestas a los datos)

Aproximación entre las distribuciones: normal, binomial y de Poisson

$N(\mu, \sigma)$

"corrección por continuidad"

$B(n, p)$ → $P(\lambda)$

1.- Si $\lambda > 10$, una $P(\lambda)$ puede aproximarse a una $N(\mu = \lambda, \sigma = \sqrt{\lambda})$