3. PRINCIPALES DISTRIBUCIONES DE PROBABILIDAD

3.1. Modelos específicos discretos. Binomial. Poisson
3.2. Modelos específicos continuos. Distribución Normal
3.3. Simulación y visualización informática. Relaciones entre modelos
3.4. Teorema central de límite
3.5. Distribuciones derivadas de la Normal.
3.6. Funciones en R

En este tema vamos a desarrollar algunos de los principales modelos de probabilidad tanto discretos como continuos. Por modelo de probabilidad "D" (o distribución-tipo "D", a menudo, simplemente "distribución D") un conjunto de distribuciones de probabilidad que tienen su función de distribución (o de cuantía/ densidad) con la misma estructura matemática habitualmente dependiente de uno o varios parámetros.

Aunque hay una enorme cantidad de distribuciones específicas de uso habitual las más importantes:

- Entre las distribuciones de variables discreta son la Distribución Uniforme de variable discreta, Distribución Binomial, la Distribución de Poisson, (en menor medida, D. Hipergeométrica).
- Entre las continuas: Distribución uniforme de variable continua, Distribución Normal, (en menor medida, exponencial, Gamma, Beta).

A las que hay que añadir un pequeño grupo de distribuciones derivadas de la Normal de aplicación en el muestreo y gran importancia en la Inferencia estadística: Distribución χ^2 (jhi-dos, Chi-dos, o Jhi-cuadrado) de Pearson, t de Student y F de Snedecor.

3.1. Modelos específicos discretos.

3.1.1. Distribución uniforme de variable discreta.

Es una distribución muy sencilla que asigna probabilidades iguales a un conjunto finito de puntos del espacio. Modeliza fenómenos en los que tenemos un conjunto de n sucesos posibles, cada uno de los cuales con la misma probabilidad de ocurrir. Si aleatorizamos de forma que cada uno de éstos sucesos se corresponda con un número natural del 1 al n obtendremos una distribución uniforme. Tendremos un único parámetro, n

Su función de cuantía definida para los valores de x = { 1, 2, … , n} vendrá dada por la constante:

P(x) = 1 /n para x = { 1, 2, … , n}

Su función de distribución vendrá dada por $F(x) = P(X \le x) = \sum_{i=1}^{x} P(x) = \sum_{i=1}^{x} \frac{1}{n} = \frac{x}{n}$

Es fácil ver que su media y su varianza son, respectivamente:

$$\mu = E(x) = \frac{n+1}{2} \qquad \sigma^2 = \frac{n^2-1}{12}$$

3.1.2. Distribución Binomial (ver en CEACES)

Es quizá la más distribución más importante de las de variable discreta. Tiene aplicación en la modelización de la distribución de la variable aleatoria :
X= "Número de resultados de cierto tipo (digamos resultados A, éxitos) en situaciones donde se repiten n pruebas dicotómicas independientes en las que P(A)=p es constante.
A este tipo de situaciones se les conoce como "procesos de bernouilli":

- Se realizan pruebas (en este caso un número determinado) separadas o separables
- Cada prueba puede dar dos resultados mutuamente excluyentes (dicotomía) digamos A y Ā (A y noA)
- La probabilidad de obtener un resultado A es constante en todas las pruebas y la llamaremos p. Obviamente la probabilidad de obtener un resultado Ā , también es constante en todas las pruebas y la llamaremos q. p+q=1

Es fácil ver que en estas circunstancias :
El recorrido de X, es {0,1,2,3,…, n} y que su función de cuantía obedece a la expresión:

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

en donde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$

es el número de combinaciones de n elementos tomados de x en x

y se corresponde con el número de formas distintas (en distintos órdenes)

en que se pueden dar x resultados A y n-x resultados no-A en n pruebas

Cuando X cumple estas condiciones estamos ante una variable que sigue una distribución binomial de parámetro n y p (n es un número natural y $0 \le p \le 1$)
Lo representamos como X→B(n,p)
A partir de aquí no es difícil obtener los valores que toman sus principales indicadores:

$$\mu = E(x) = \sum_{x=1}^{n} x.P(x) = np$$

$$\sigma^2 = \text{var}(x) = E\left[\left(x - E(x)\right)\right]^2 = npq \qquad \text{(máxima para p=q=0.5)}$$

$$Moda(x) = mo/: \quad np - q \le mo \le np + p \quad \text{(puede haber una o dos)}$$

$$\text{asimetría: } \gamma_1 = \frac{q-p}{\sqrt{npq}} \qquad \text{(Simétrica para p=q=0.5)}$$

$$\text{curtósis : } \gamma_2 = \frac{1-6pq}{npq} \qquad \text{(mínima para p=q=0.5)}$$

La forma más fácil de calcular media, varianza, asimetría y curtosis es a partir de su FGM , derivándola sucesivamente en el punto t=0 y obteniendo así $\alpha_r = E(x^r) = \varphi^{(r)}(t=0)$.

Y su FGM es :

$$\varphi(t) = E\left[e^{tx}\right] = \sum_{x=1}^{n} e^{tx} \cdot P(x) = \sum_{x=1}^{n} \binom{n}{x} e^{tx} p^x q^{n-x} = \sum_{x=1}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x} = (pe^t - q)^n$$

Cuando n=1 (sólo hay una prueba estamos ante lo que se conoce como distribución dicotómica, binaria o de Bernouilli.

Si p y q no son constantes (las pruebas no son indepependientes; si se trata de extracciones no hay remplazamiento estamos ante lo que se conoce como Distribución Hipergeométrica (ver cuadro resumen, más abajo)

Visualización de la Binomial : https://www.desmos.com/calculator/csmbp1smwr

3.1.3. Distribución de Poisson (ver en CEACES)

Esta distribución es una de las más importantes distribuciones de variable discreta. Sus principales aplicaciones hacen referencia a la modelización de situaciones en las que nos interesa determinar el número de hechos de cierto tipo que se pueden producir en un intervalo de tiempo o de espacio, bajo presupuestos de aleatoriedad y ciertas circunstancias restrictivas. Otro de sus usos frecuentes es la consideración límite de procesos dicotómicos reiterados un gran número de veces si la probabilidad de obtener un éxito es muy pequeña (convergencia Binomial / Poisson)

Esta distribución se puede hacer derivar de un proceso experimental de observación en el que tengamos las siguientes características

· Se observa la realización de hechos de cierto tipo durante un cierto periodo de tiempo o a lo largo de un espacio de observación

· Los hechos a observar tienen naturaleza aleatoria; pueden producirse o no de una manera no determinística.

· La probabilidad de que se produzcan un número x de éxitos en un intervalo de amplitud t no depende del origen del intervalo (Aunque, sí de su amplitud)

· La probabilidad de que ocurra un hecho en un intervalo infinitésimo es prácticamente proporcional a la amplitud del intervalo.

· La probabilidad de que se produzcan 2 o más hechos en un intervalo infinitésimo es un infinitésimo de orden superior a dos.

En consecuencia, en un intervalo infinitésimo podrán producirse 0 ó 1 hecho pero nunca más de uno

· Si en estas circunstancias aleatorizamos de forma que la variable aleatoria X signifique o designe el "*número de hechos que se producen en un intervalo de tiempo o de espacio*", la variable X se distribuye con una distribución de parámetro λ.

El parámetro de la distribución es, en principio, el factor de proporcionalidad para la probabilidad de un hecho en un intervalo infinitésimo. Se le suele designar como parámetro de intensidad, aunque se corresponde con el número medio de hechos que cabe esperar que se produzcan en un intervalo unitario (media de la distribución); y que también coincide con la varianza de la distribución.

Por otro lado es evidente que se trata de un modelo discreto y que el campo de variación de la variable será el conjunto de los número naturales, incluido el cero: $x \in \{0,1,2,3,\ldots\}$

Función de cuantía: $P(x)=P(X=x)=\frac{e^{-\lambda}\lambda^{x}}{x!}$

Función de distribución: $F(x)=P(X\le x)=\sum_{i=0}^{x}\frac{e^{-\lambda}\lambda^{i}}{i!}$

Función generatriz de momentos: $\varphi(t)=e^{\lambda(e^{t}-1)}$

Media: $\mu=E(x)=\sum_{x=0}^{\infty}x\frac{e^{-\lambda}\lambda x}{x!}=\varphi'(t=0)=\lambda$ Varianza: $\sigma^2=E\left[(x-\mu)^2\right]=\lambda$

Moda: $\lambda-1\le \text{moda}\le\lambda$

Visualización: https://www.desmos.com/calculator/xowiefjxg5

RESUMEN DE LAS DISTRIBUCIONES DE V. DISCRETA MÁS IMPORTANTES.

Distribución	**Situación/ aleatorización**	**F.cuantía y rango**	**Media**	**Varianza**
Dicótomica X → D(p)	Número de éxitos en una prueba con p(éxito)=p	$P(x)=p^{x}q^{1-x}\quad x\in\{0,1\}$	p	pq
Binomial X→ B(n,p)	Número de éxitos en n pruebas con p constante	$P(x)=\binom{n}{x}p^{x}q^{n-x}$ $x\in\{0,1,\dots,n\}$	np	npq
Geométrica X→G(p)	Número de pruebas necesarias para obtener el primer éxito	$P(x)=q^{x-1}p\quad x\in\{1,2,\dots,\infty\}$	$\frac{1}{p}$	$\frac{q}{p^2}$
Binomial negativa X→BN(k,p)	Número de pruebas necesarias para obtener los k primeros éxitos	$P(x)=\binom{x-1}{k-1}p^{k}q^{x-k}$ $x\in\{k,k+1,\dots,\infty\}$	$\frac{k}{p}$	$\frac{kq}{p^2}$
Hipergeométrica X→H(N,n,p)	Número de éxitos en n pruebas de extracción sin reposición sobre una población de tamaño N	$P(x)=\frac{\binom{Np}{x}\cdot\binom{Nq}{n-x}}{\binom{N}{n}}\quad x\in\{0,1,\dots,\min(n,Np)\}$	np	$\frac{N-n}{N-1}pq$
Poisson X→P(λ)	Número de hechos que ocurren en un intervalo (espacio) unitario supuesto homogeneidad de t e imposibilidad de simultaneidad	$P(x)=\frac{e^{-\lambda}\lambda^{x}}{x!}\quad x\in\{0,1,..\infty\}$	λ	λ
Uniforme de variable discreta	Numeral del individuo seleccionado de entre n equiprobables	$P(x)=\frac{1}{n}\quad x\in\{1,2,\dots n\}$	$\frac{n+1}{2}$	$\frac{(n^2-1)}{12}$

Para muchas de estas distribuciones podemos calcular sus probabilidades con la aplicación libre CaEst o través del entorno R o R-Commander dónde también podemos obtener muestras y representaciones gráficas.

En esta tabla aparecen los comandos en código con los que podemos manipular en R algunas de esta distribuciones

Distribución	F. cuantía	F. Distribución	Inversa de F(x) cuantil	Generación de una muestra
Binomial	dbinom(x,n,p)	pbinom(x,n,p)	qbinom (prob,n,p)	rbinom (nn,n,p)
B.Negativa	dnbinom(x,k,p)	pbinom(x,k,p)	qbinom (prob,k,p)	rbinom (nn,k,p)
Poisson	dpois(x,λ)	ppois(x, λ)	qpois (prob, λ)	rpois (nn, λ)
Hipergeom.	dhyper(x,Np,Nq,n)	phyper(x,Np,Nq,n)	qhyper(pr,Np,Nq,n)	rhyper(nn, Np,Nq,n)
Geométrica	dgeom(x,p)	pgeom(x,p)	qgeom(prob,p)	rgeom(nn,p)

(Otras muchas distribuciones también son susceptibles de ser tratadas con commandos similares dnombre(x, parámetros),pnombre(x, parámetros), qnombre(prob,parámetros),rnombre (nn,parámetros) ,donde nombre es una abreviatura del nombre de la distribución)

Ejemplo:

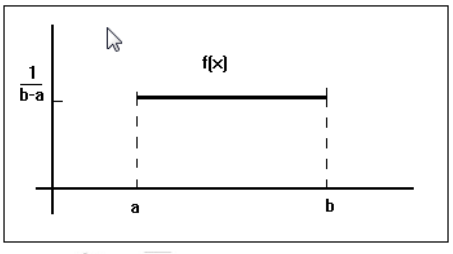
```
#valor que cubre un 90% de probabilidad en una BN(k=3, p=0.3)
qnbinom(0.9,3,0.3)

#grafica función de cuantía(?) de una B(n=10,p=2)
x<- seq (0,10)
y<-dbinom(x,10,0.2)
plot(x, y, type = "h", xlab = "x", ylab = "P(X)", xlim = c(0,
                                        + 10), ylim =c(0,0.5) )
#grafica de la F. distribucion de una P(lambda=5)
y1<-ppois(x,5)
x<- seq (0,10)
plot(x, y1, type = "l", xlab = "x", ylab = "F(X)", xlim = c(0,
                                      + 10), ylim =c(0,1) )
#muestra aleatoria de 100 datos procedentes de una H(N=52,n=5,p=0.25) con histogramma
muestra<-rhyper(100,13,39,5)
hist(muestra)
```

3.1. Modelos específicos continuos.

3.1.1.Distribución uniforme de variable continua.

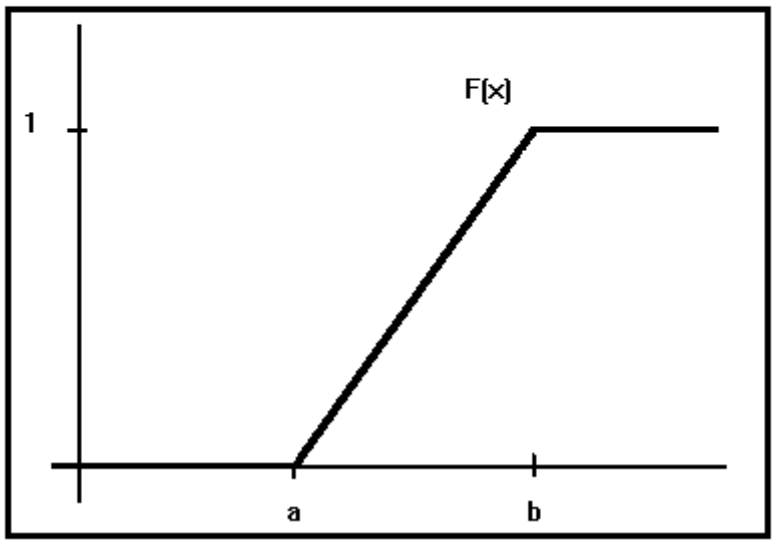
Como su nombre indica el planteamiento de la distribución es el hecho de que la probabilidad se distribuye uniformemente a lo largo de un intervalo . Así : dada una variable aleatoria continua, X , definida en el intervalo [a,b] de la recta real, diremos que x tiene una distribución uniforme en el intervalo [a,b] cuando su función de densidad para sea:

$$f(x) = \frac{1}{b-a} \qquad \text{para } x \in [a,b]$$



Y la función de distribución :

$$F(X) = P(x \le X) = \begin{cases} 0 \text{ si } X < a \\ \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a} \quad \text{si } a \le X \le b \\ 1 \text{ si } X \ge b \end{cases}$$



La media es, obviamente el valor central del intervalo:

$$\mu = E(X) = \int_a^b \frac{x-a}{b-a} dx = \frac{(b+a)}{2}$$

Y la varianza : $\sigma^2 = \frac{(b-a)^2}{12}$

```
#distribucion uniforme
runif(5,0,10)#muestra de tamaño 5 de una uniforme [0,10]
dunif(1,0,10);dunif(3.5,0,10)#f. de dendidad (f=1/10=0.1)
punif(7.3,0,20)# función de distribución de 7.3 de  U[0,20]
qunif(0.75,10,50)#percentil 75 de U[10,50]
```

3.1.2.Distribución exponencial

La distribución exponencial sirve para modelizar el lapso de tiempo entre dos eventos consecutivos de Poisson que ocurren de manera independiente y a una frecuencia constante. Esta distribución se emplea para problemas de tiempo, de tiempo de fallo, líneas de espera, supervivencia etc. En otras palabras, en un proceso de Poisson donde se repite sucesivamente un experimento a intervalos iguales de tiempo, el tiempo que transcurre entre la ocurrencia de dos sucesos consecutivos sigue un modelo probabilístico exponencial.
Función de densidad

$$f(x)=\begin{cases}\alpha e^{-\alpha x} & x\geq 0\\ 0 \quad \text{en los demás casos}\end{cases}$$

E(X)=1/α y Var (X)=1/α^2

```
#distribución exponencial
rexp(3,5)# muestra de tamaño 3 de una exp(alfa=5)
dexp(2,10)#f. de densidad en 2 de una exp(alfa=10)
pexp(2,0.5)#F. de distribución en 2 exp(alfa=0.5)
qexp(0.5,0.03)#mediana de exp(alfa=0.03)
curve(dexp(x, 0.03), 0, 40)#gráfica de la densidad exp(0.03)
```

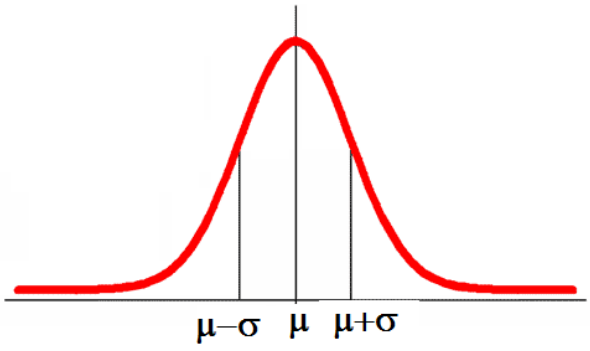
3.2.3.Distribución normal

Decimos que una variable aleatoria continua X se distribuye como una normal, o que se distribuye normalmente con parámetros (μ, σ), lo que representamos como:
X→ N(μ, σ) cuando su función de densidad.

$$\mathbf{f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}} \qquad x\in]-\infty,\infty[$$

Se comprueba que la esperanza matemática y la varianza son respectivamente los parámetros μ y σ^2 (σ es la desviación típica) de la función de densidad.

La representación gráfica de esta función de densidad f(x), da lugar a una curva campaniforme perfecta con máximo en el punto de abscisa x= μ y el punto de inflexión sobre μ ±σ. La curva es además simétrica respecto del eje de abscisa en el punto x= μ y tiene por asíntota horizontal ese mismo eje de abscisas.



Decimos que una variable aleatoria Z absolutamente continua se distribuye como una normal, o que se distribuye normalmente con parámetros (0,1) cuando su función de densidad sea la de una normal con μ=0 y σ=1 y se denota Z→N (0,1).
La función de densidad será:

$$f(z)=\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} \qquad x\in]-\infty,\infty[$$

La distribución normal conserva la linealidad en todas sus formas.

1) Si $X \rightarrow N(\mu_x, \sigma_x)$
entonces una transformación lineal de X, digamos $Y=a+bX$,
también será una variable aleatoria normal con $\mu_y=a+b\mu_x$ y $\sigma_y=b.\sigma_x$
esto es,

$$Y \rightarrow N(\mu_y=a+b\mu_x \; ; \; \sigma_y=b.\sigma_x)$$

La tipificación y su transformación inversa son casos particulares de esto.

2) Teorema de adición:
La suma de variables normales **independientes** es también una variable aleatoria normal con media, la suma de las medias y **varianza**, la suma de las **varianzas**.

Sean X e Y dos variables aleatorias independientes y tales que:

$$X \rightarrow N[\mu_x \; ; \; \sigma_x]$$

$$Y \rightarrow N[\mu_y \; ; \; \sigma_y]$$

queremos comprobar que la variable $U=X+Y$ es tal que:

$$U \rightarrow N\left[\mu_x + \mu_y ; \sqrt{\sigma_x^2 + \sigma_y^2}\right]$$

Resultado trivialmente extrapolable a cualquier número de variables sumando.

Ejemplo: Dos variables X e Y son independientes y , además se distribuyen normalmente de manera que: $X \rightarrow N(2;1)$
mientras que $Y \rightarrow N(4;2)$; si $Z=X+Y$

la distribución y parámetros de la variable Z será: $Z \rightarrow N(6 \; ; \; 2{,}236)$

$$\mu_z = \mu_x + \mu_y = 2 + 4 = 6$$

$$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2} = \sqrt{1^2 + 2^2} = \sqrt{5}$$

3) Teorema fundamental:

Cualquier combinación lineal de variables aleatorias normales **independientes**
Da como resultado una nueva variable normal
cuya **media** será la misma **combinación lineal** de las medias
y cuya **varianza** será la combinación lineal de las varianzas **con los términos** elevados al **cuadrado** y **sin** término independiente, esto es:

Sean las variables aleatorias X_i , con i=1,2,3,...n, todas ellas independientes tales que:

$x_i \Rightarrow N(\mu_i; \sigma_i)$ y sean los números reales **a ¡** con i=0,1,2,3,...n

la variable combinación lineal : $Y = a_0 + a_1X_1 + a_2X_2 + ... + a_nX_n = a_0 + \sum_{i=1}^{n} a_iX_i$

se distribuirá según : $Y \to N\left[a_0 + \sum_{i=1}^{n} a_i\mu_i \; ; \sqrt{\sum_{i=1}^{n} a_i^2\sigma_i^2} \right]$

4) Linealidad en el caso más general.
Cualquier combinación lineal de variables aleatorias normales aunque no sean independientes sigue también una distribución normal con media, la misma **combinación lineal** de las medias y con varianza, la combinación lineal de las varianzas **con los términos** elevados al **cuadrado** y **sin** término independiente a la que hay que añadir el término de los dobles productos de los coeficientes y covarianzas correspondientes.
Es decir :
Si tenemos n variables aleatorias nor*males*
$x_i \Rightarrow N(\mu_i; \sigma_i)$
pero no independientes de modo que :
$\sigma_{ij} = \text{cov}(X_i, X_j) = \rho_{ij}\sigma_i\sigma_j$ son las correspondientes covarianzas entre X_i e X_j
entonces

la variable combinación lineal : $Y = a_0 + a_1X_1 + a_2X_2 + ... + a_nX_n = a_0 + \sum_{i=1}^{n} a_iX_i$

se distribuirá según : $Y \to N\left[a_0 + \sum_{i=1}^{n} a_i\mu_i \; ; \sqrt{\sum_{i=1}^{n} a_i^2\sigma_i^2 + 2\sum_{i \neq j} a_ia_j\sigma_{ij}} \right]$

En virtud del **TCL y otros teoremas límite** la distribución normal es la distribución límite de otras en un gran número de ocasiones.

3.4. Teorema central del límite. Teoremas límite. Convergencia

Tanto el teorema central del límite (TCL), en sus dos formulaciones más generales (Lyounov y Lindenberg-Levy) como otros teoremas límite de gran importancia práctica se deben enunciar y estudiar adecuadamente en el marco de la llamada teoría de la convergencia de sucesiones de variables aleatorias. Sin embargo, aquí nos preocuparemos fundamental de sus condiciones de aplicación, de su trascendencia de cara a la inferencia estadística y de su aplicabilidad práctica a distintos contextos.

Antes de ocuparnos del TCL le dedicaremos alguna atención a otros teoremas de convergencia de importancia: La convergencia binomial/poisson, el teorema de Moivre (convergencia binomial / normal), la convergencia Poisson/ Normal.

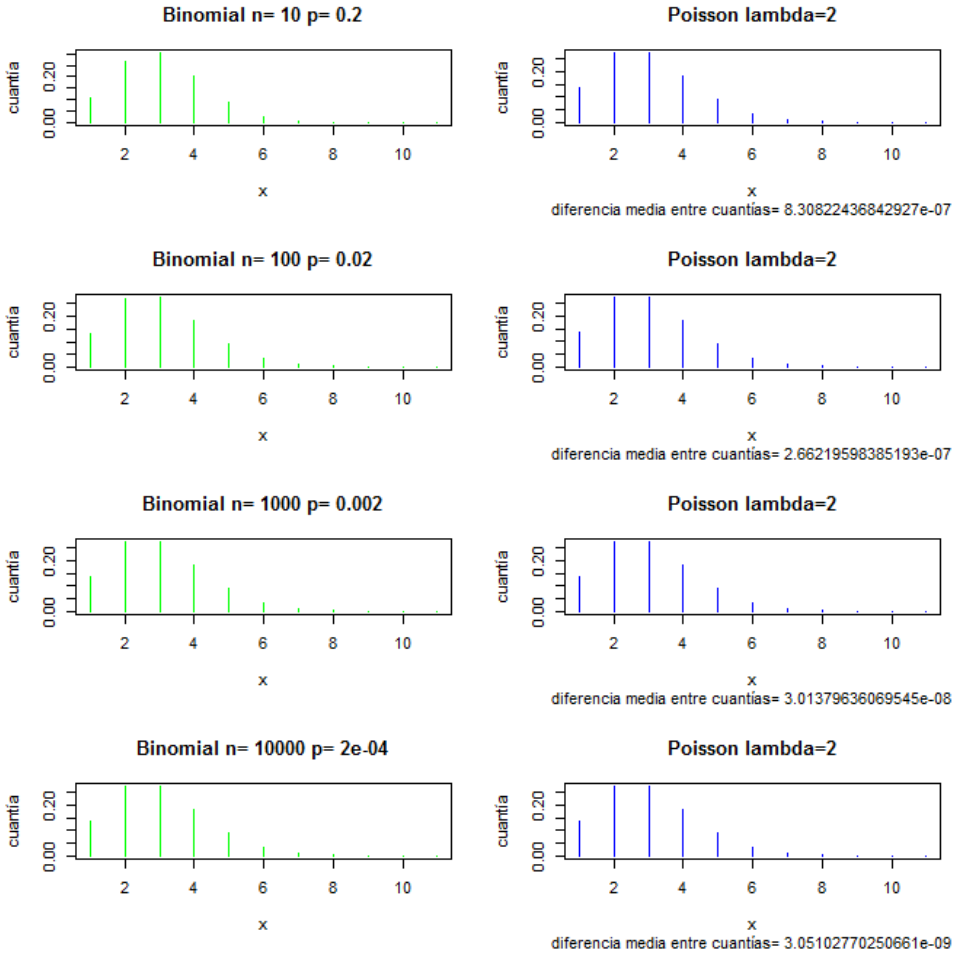
Convergencia Binomial/ Poisson.

Una variable aleatoria binomial (n,p) tiende a ir asemejándose en su distribución a una variable Poisson con $\lambda = np$ conforme el parámetro n va creciendo hacia infinito a la vez que el parámetro p decrece hacia cero (a la misma velocidad: el producto np es de una magnitud finita moderada).
Esta tendencia a asemejarse la entendemos como que su función de cuantía va convergiendo cada vez más hacia la f. de cuantía de una distribución de Poisson:

$$\lim_{\substack{n\to\infty \\ p\to 0}} P(x \mid X \to B(n,p) = \lim_{\substack{n\to\infty \\ p\to 0}} \binom{n}{x} p^x (1-p)^{(n-x)} = \frac{e^{np}.(np)^x}{x!} = P(x \mid X \to Pois(\lambda = np))$$

Este resultado permite, en la práctica resolver problemas relativos a variables binomiales con n grande y p pequeña como si fueran variables de Poisson.
Visualización con R :

```
#convergencia binomial-poisson
n<-c(10,100,1000,10000)
p<-c(0.2,0.02,0.002,0.0002)
windows()
par(mfrow=c(4,2))
for(i in 1:length(n)){
   dif=sum(dbinom(0:10,n[i],p[i])-dpois(0:10,lambda=n[i]*p[i]))/10
   plot(dbinom(0:10,n[i],p[i]),type="h", col="green",
                  xlab= "x",ylab="cuantía" ,main=paste("Binomial n=",n[i],"p=",p[i]))
   plot(dpois(0:10,lambda=n[i]*p[i]),type="h",col="blue",main="Poisson lambda=2",
                  xlab="x",ylab="cuantía", sub= paste("diferencia media entre cuantías=",dif))
 }
```



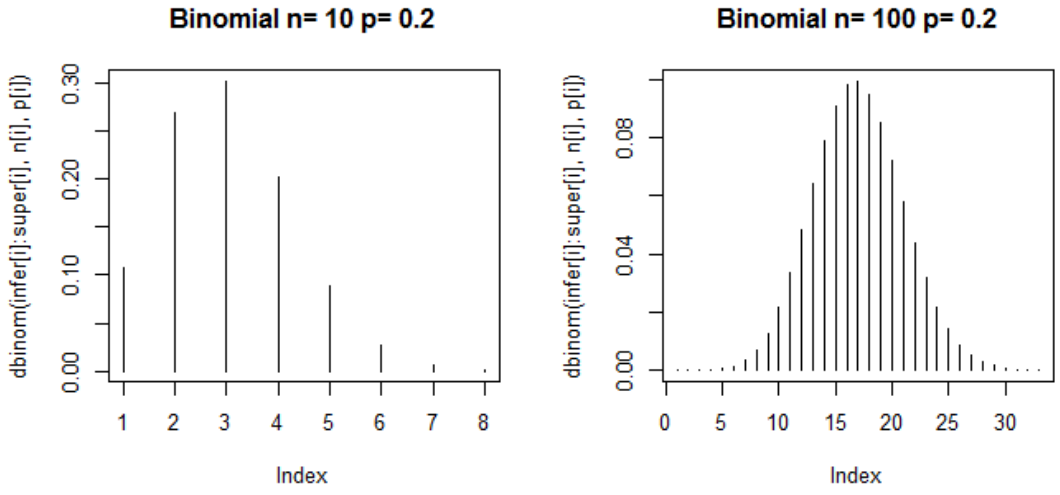
Teorema de Moivre

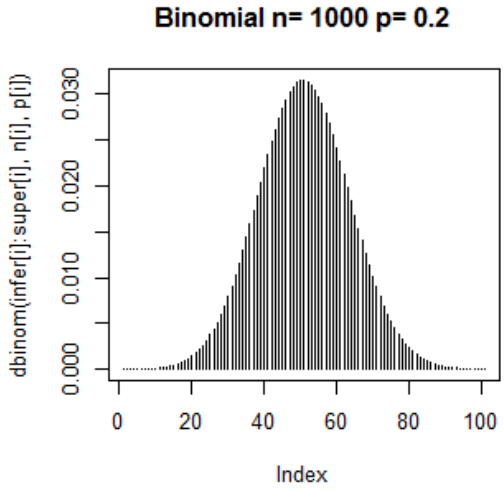
El Teorema de Moivre garantiza que una sucesión de variables aleatorias que siguen distribuciones binomiales con parámetro n cada vez más grande de forma creciente converge a una variable aleatoria Normal con media, np y varianza, npq.

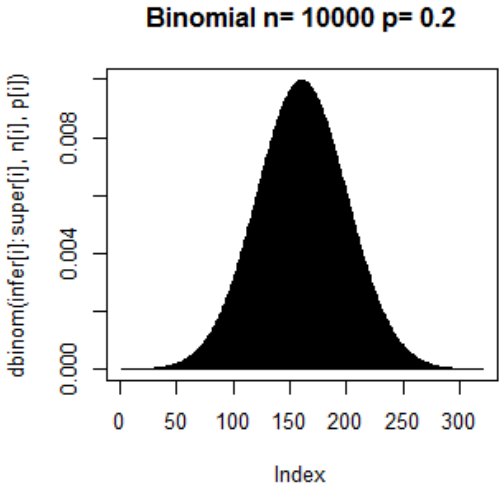
$\{X_n\}$ $X \rightarrow B(n,p)$ converge cuando $n \rightarrow \infty$ a Y / $Y \rightarrow N(\mu=np\ ;\ \sigma=(npq)^{1/2})$

Visualización con R :

```
#teorema de Moivre
n<-c(10,100,1000,10000)
p<-c(0.2,0.2,0.2,0.2)
 infer<-matrix(NA,nrow = 1,ncol=length(n))
 super<-matrix(NA,nrow = 1,ncol=length(n))
for(i in 1:length(n)){
  infer[i]<-max(0,floor(n[i]*p[i]-floor(4*sqrt(n[i]*p[i]*(1-p[i])))))
  super[i]<-floor(n[i]*p[i]+floor(4*sqrt(n[i]*p[i]*(1-p[i]))))
 }
 windows()
 par(mfrow=c(2,2))
 for(i in 1:length(n)){
    plot(dbinom(infer[i]:super[i],n[i],p[i]),type="h",xlab=NULL,ylab=NULL,
            main=paste("Binomial n=",n[i],"p=",p[i]) )
}
```





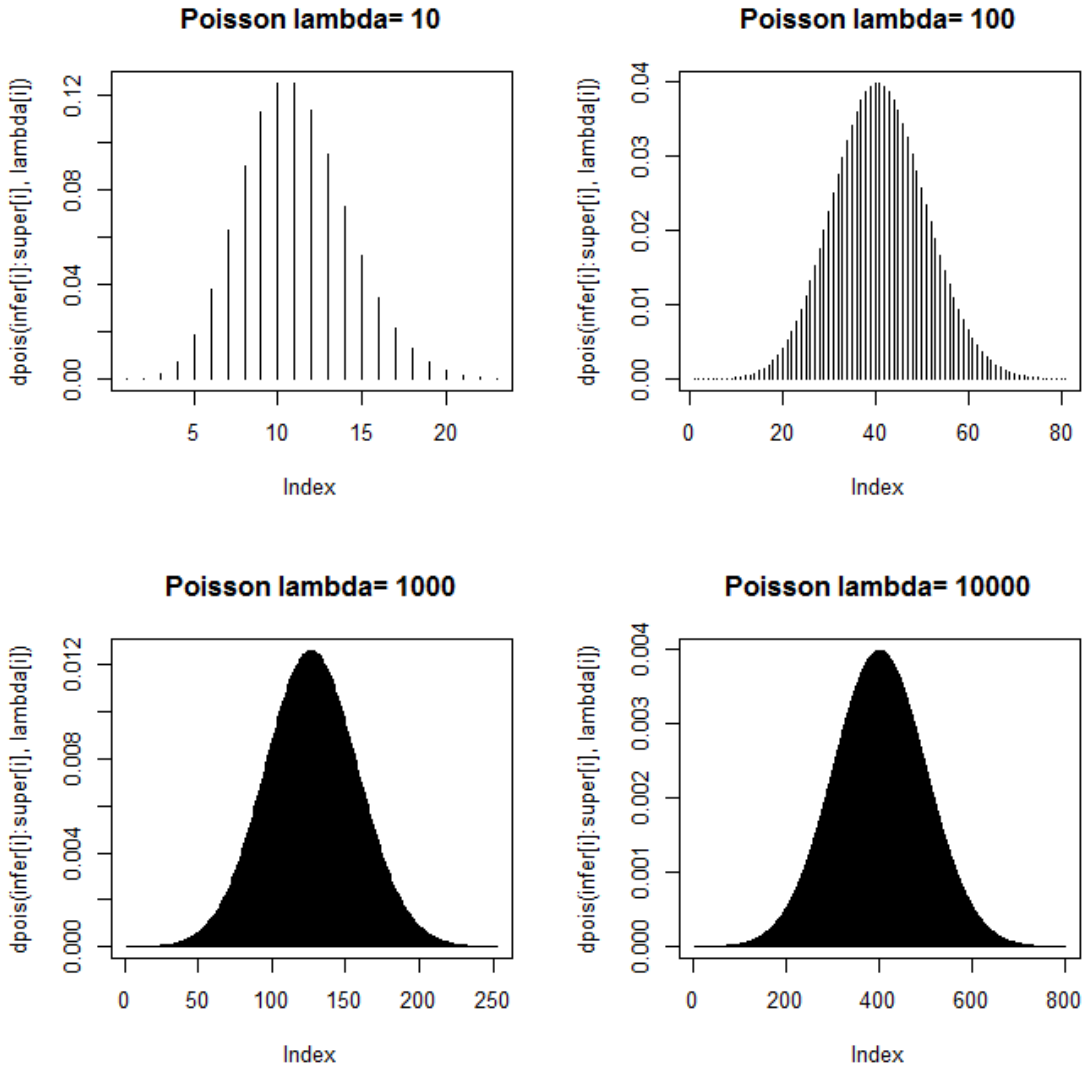


Convergencia Poisson/ Normal

De igual forma, la distribución convergen también a distribución normal conforme su parámetro λ tiende a infinito.

Visualización con R:

```
#CONVERGENCIA pOISSON /nORMAL
lambda<-c(10,100,1000,10000)
 infer<-matrix(NA,nrow = 1,ncol=length(n))
super<-matrix(NA,nrow = 1,ncol=length(n))
for(i in 1:length(lambda)){

 infer[i]<-max(0,lambda[i]-floor(4*sqrt(lambda[i])))
 super[i]<-lambda[i]+floor(4*sqrt(lambda[i]))
}
windows()
par(mfrow=c(2,2))
for(i in 1:length(n)){
 plot(dpois(infer[i]:super[i],lambda[i]),type="h",xlab=NULL,ylab=NULL,
     main=paste("Poisson lambda=",lambda[i]  ))
}
```



Vemos como en un sentido muy amplio parece que casi cualquier disrtibución cuando sus parámetros crecen empieza aproximarse cada vez más a una distribución normal y, efectivamente es así bajo condiciones bastante generales que son las que se consignan en los llamados TCL.

TCL (forma de Lindeberg-Levy):

La suma de un gran número (n) de variables aleatorias independientes todas ellas con la misma distribución de igual media e igual varianza tiende a distribuirse con una distribución normal con media, n veces la media común y varianza, n veces la varianza común.

$X_1, X_2, ..., X_n$ todas independientes dos a dos y $X_i \to Distrib.[\mu, \sigma]$

entonces si $n \to \infty$: la variable $S=\sum_{i=1}^{n} X_i$ | $S \to N[n\mu ; \sigma_s = \sqrt{n\sigma^2}]$

$$S \to N[n\mu ; \sigma_s = \sigma\sqrt{n}]$$

Obviamente si a esta variable suma (S) la dividimos por el número de sumandos el resultado también será una variable normal, por la linealidad de la distribución normal.

Y este cociente no es otra cosa que la media aritmética de esas n variables aleatorias.

$X_1, X_2, ..., X_n$ todas independientes dos a dos y $X_i \to Distrib.[\mu, \sigma]$

entonces si $n \to \infty$: la variable $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{S}{n}$ | $\bar{X}_n \to N[\mu ; \sigma_{\bar{X}_n} = \frac{\sigma}{\sqrt{n}}]$

la normalidad está garantizada por la linealidad y la media y desviación típica resultan de:

$$E[\bar{X}_n] = E[\frac{S}{n}] = \frac{E[S]}{n} = \frac{n\mu}{n} = \mu$$

$$\text{var}[\bar{X}_n] = \text{var}[\frac{S}{n}] = \frac{\text{var}[S]}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \Rightarrow DT[\bar{X}_n] = \sqrt{\text{var}[\bar{X}_n]} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

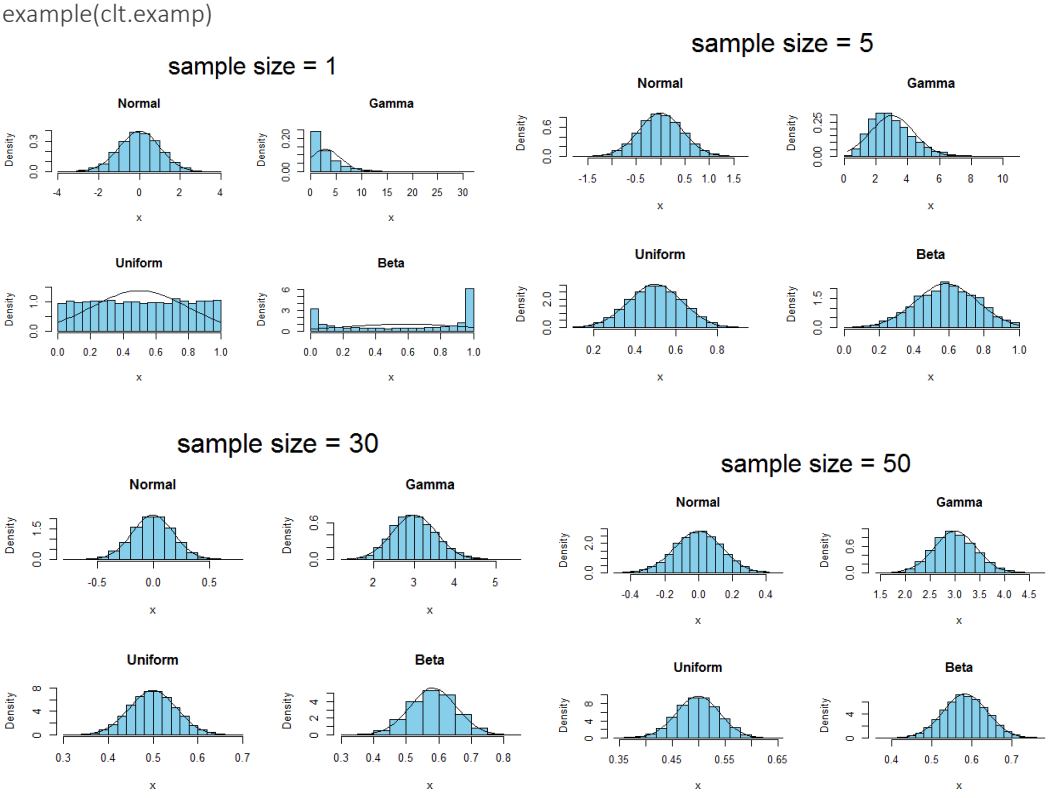
La transcendencia del TCL en el muestreo y la inferencia viene del hecho de que si realizamos una muestra aleatoria (con reposición) de una población, cada dato extraído es una v.a. cuya distribución es la de la población cuya esperanza es la media de la población y cuya varianza es la de la población (para todos los datos).

Si la muestra es con reposición cada dato es independiente de todos los demás y por lo tanto se cumplen las condiciones del TCL.:

La media de una muestra aleatoria (con remplazamiento) tenderá a distribuirse con una Normal si la muestra es lo suficientemente grande.

Visualización en R :

```
Install.packages("TeachingDemos")
library(TeachingDemos)
example(clt.examp)
```



3.5.Distribuciones derivadas de la normal

A partir de determinadas transformaciones de variables normales se generan algunas variables cuyas distribuciones son de vital importancia en la inferencia sobre poblaciones que o bien son normales o bien son aproximadamente normales. Entre estas están la distribución χ^2 (Chi-dos Jhi-dos, Jhi-cuadrado, Chi-cuadrado) de Pearson, la distribución t de Student y la distribución F de Snedecor.

χ^2 de Pearson (Chi-dos):

A partir de n variables normales[0,1] independientes: Z1,Z2,…,Zn si las elevamos al cuadrado y las sumamos obtendremos una variable aleatoria cuya distribución llamaremos distribución χ^2 de Pearson con n grados de libertad. Y→ χ^2_n

En ceaces: http://www.uv.es/ceaces/normaMu/chi2/chi2.htm

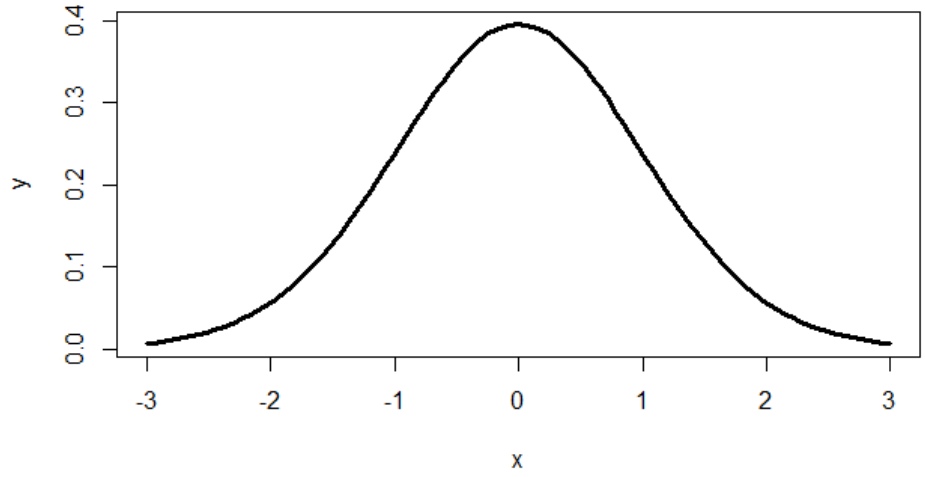
```
#chi-dos
ind <- c(1, 2)
for (i in ind){
  curve(dchisq(x, df = i), 0, 10, add = TRUE)
}

ind <- c(3, 4,10, 15)
for (i in ind){
  curve(dchisq(x, df = i), 0, 10, add = TRUE)
}
```

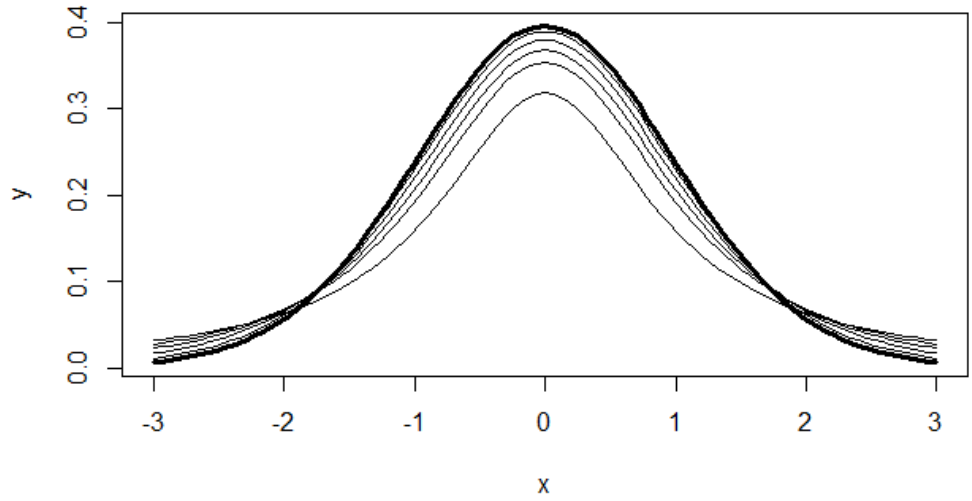
Distribución t de student.:

Dada una variable aleatoria z que siga una N[0,1] y otra variable Y→ χ^2_n independiente de la anterior:

El cociente $t = \frac{Z}{\sqrt{\frac{Y}{n}}}$ será una nueva variable aleatoria que seguirá una t de Student con n grados de libertad ($\mathbf{t}_n$)

```
#t de student
curve(dt(x, df = 30), from = -3, to = 3, lwd = 3, ylab = "y")
ind <- c(1, 2, 3, 5, 10)
for (i in ind){
  curve(dt(x, df = i), -3, 3, add = TRUE)
}
```





En CEACES: http://www.uv.es/ceaces/normaMu/t/t.htm

F de Snedecor:

Dadas dos variables independientes X e Y ambas con distribuciones χ^2 el cociente entre ambas conveniente divididas por sus grados de libertad será una nueva variable aleatoria cuya distribución será una F .

$X \to \chi^2_m$ $\quad$ $Y \to \chi^2_n$ $\quad$ X e Y independientes

$$F = \frac{X/m}{Y/n} \to F_{m,n}$$

Queda claro por tanto que la distribución F de Snedecor tiene dos parámetros , que son m y n ; grados de libertad del numerador , grados de libertad del denominador.

En CEACES: http://www.uv.es/ceaces/normaMu/f/f.htm

```
#F de snedecor
m <- c(4, 15)
n <- c(10, 5)
for (i in m){
  for( j in n){
  curve(df(x, i,j), 0, 10, add = TRUE)
}}
```