

1 Teorema central de límite

Sean X_1, X_2, \dots, X_k variables aleatorias independientes e idénticamente distribuidas con media μ y varianza $\sigma^2 < \infty$, entonces

$$\frac{\sum_{i=1}^k X_i - k\mu}{\sqrt{k\sigma^2}} \xrightarrow{d} Z$$

donde $Z \sim N(0,1)$

2 Ejercicio

Utiliza el Teorema para generar un programa que simule una variable aleatoria $N(0,1)$ a partir de *binomiales*(n,p).

Generen seis histogramas de 10000 v.a. normales con los siguientes parámetros:

- $n = 10, p = 0.5, k = 10, 20, 100, 1000$
- $n = 10, p = 0.99, k = 10, 20, 100, 1000$

Comparen los histogramas con la función de densidad teórica de la normal estándar, para esto pueden hacer:

```
sop = seq(-6,6, length.out = 1000)
par(new = TRUE )
lines(sop, dnorm(sop))
```

El comando dará error si no has generado antes el histograma.

Observaciones: Para generar k *Binomiales*(n,p) en R existe el comando *rbinom*(k,n,p), la media de una v.a. *Binomial*(n,p) es $n \cdot p$ y su varianza es $np(1-p)$. Su reporte debe tener ocho figuras.

¿Cuál converge más rápido? ¿Por qué?

Teorema central de límite

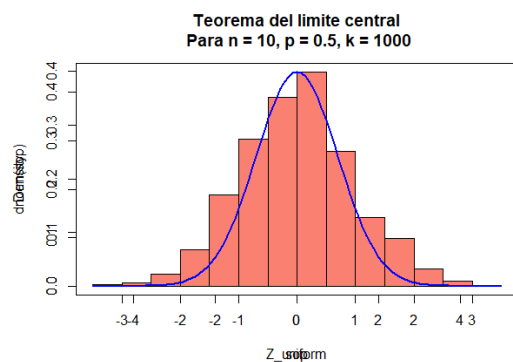
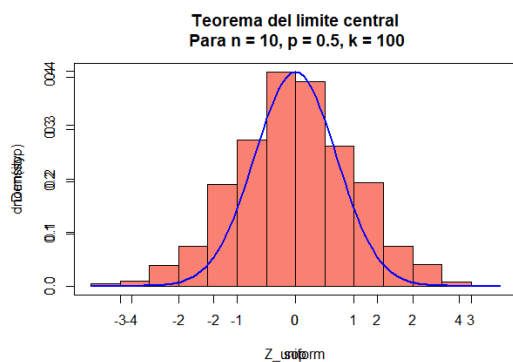
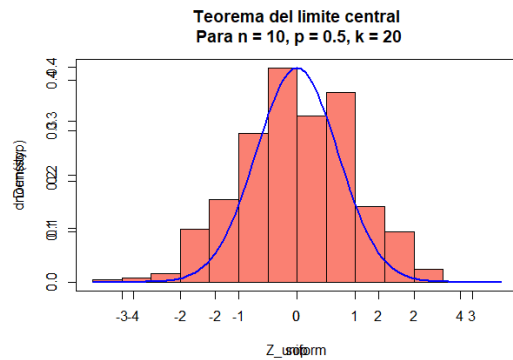
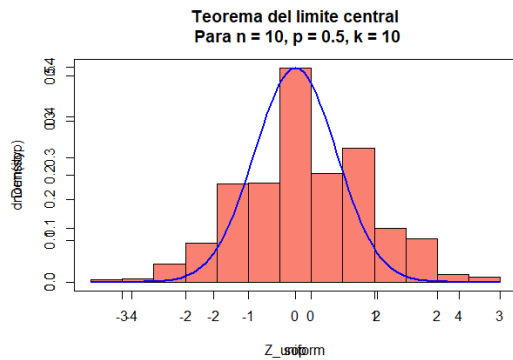
Para implementar el teorema del limite central, crearemos una funcion que se encargara de encontrar la variables aleatoria z , estará formada de esta forma:

```
1 #' Teorema del limite central
2 #'
3 #' @param X - vector de variables aleatorias
4 #' @param mu - media
5 #' @param varianza - varianza
6 #'
7 #' @return - variable aleatoria N(0,1)
8 teorema_limite_central <- function(X, mu, varianza) {
9   k <- length(X)
10   z <- (sum(X) - k*mu) / (k*varianza)^(1/2)
11   return(z)
12 }
```

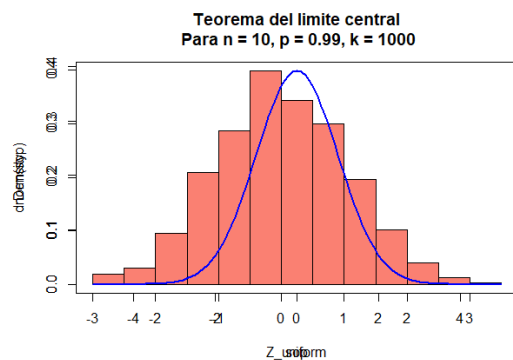
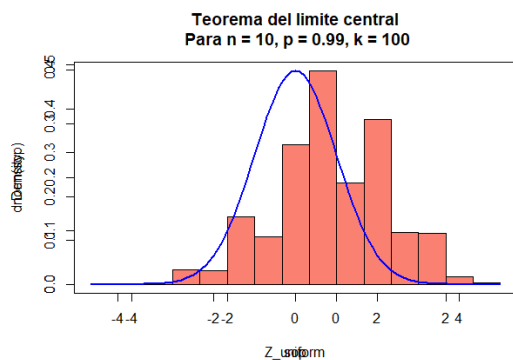
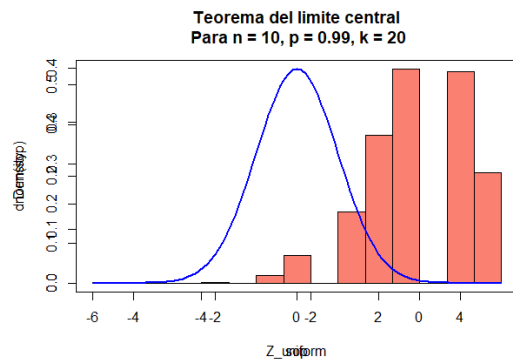
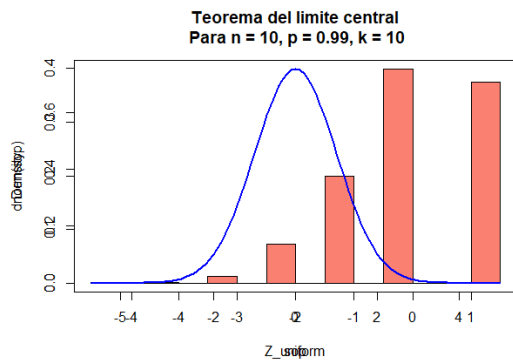
Para encontrar la convergencia hacia una distribución normal, lo que haremos será generar 1000 variables aleatorias usando el teorema del límite central, esto nos dará variables aleatorias que corresponden a una distribución normal $N(0,1)$, para demostrarlo implementaremos esta función:

```
1  #' Simula variable aleatorias N(0,1) a partir de binomiales(n,p)
2  #'
3  #' @param N - numero de variables aleatorias Z
4  #' @param n - numero de ensayos para la
5  #' @param p - Probabilidad de éxito
6  #' @param k - total de binomiales
7  #'
8  #' @return - una grafica y un histograma
9  simular_Z <- function(N, n, p, k, msg="") {
10    q <- 1-p
11    mu <- n*p
12    varianza <- n*p*q
13    Z_uniform <- c()
14    for (i in 1:N){
15      binomiales <- rbinom(k, n, p)
16      Z_uniform[i] <- teorema_limite_central(binomiales, mu, varianza)
17    }
18    hist(Z_uniform, main=paste('Teorema del limite central\n',msg),
19         col='salmon', prob = TRUE)
20    par(new = TRUE )
21    sop <- seq( -5 ,5 , by=0.1 )
22    plot(sop, dnorm(sop), type='l', col='blue', lwd=2)
23  }
```

En esta función generaremos 1000 variables aleatorias uniformes y las graficaremos para demostrar nuestra hipótesis, graficaremos un histograma de nuestras variables Z y la vez la silueta de la distribución normal estándar. Ahora ya no serán promedios parciales como vimos anteriormente para la ley de los grandes números, sino usaremos suma parciales renormalizadas en cada iteración. Un detalle importante es que generaremos variables binomiales que convergerán a la distribución normal.



Vemos que cuando tenemos como parámetro $p = 0.5$ nuestro histograma fácilmente se puede observar como converge a una distribución normal, a pesar haber variado los parámetros $k = 10, k = 20, k = 100, k = 1000$ no varia mucho, sigue teniendo la silueta de una normal. Puede deberse que le hemos asignado una probabilidad fija de 0.5.



Podemos ver un gran contraste cuando tenemos $P = 0.99$ para k diferente, al principio con una $k = 10$ los valores están muy alejados hacia los costados, conforme vamos aumentando k el histograma se va comportando más a la distribución normal, moviendo haciendo a la derecha y llenando poco a poco la silueta de la distribución normal.

Podemos concluir que dada cualquier función de distribución y usando el teorema del límite central, las variables de esa distribución convergerá a hacia una distribución normal estándar. Esto es increíble dado que reafirmamos nuevamente la aparición de la distribución normal y su importancia, ya que para cualquier variable aleatoria de una distribución, usando el TLC veremos convergerá a una normal.