

Unidad III

Parte III

Estimación de parámetros:

Estimación puntual

Estimación por intervalos

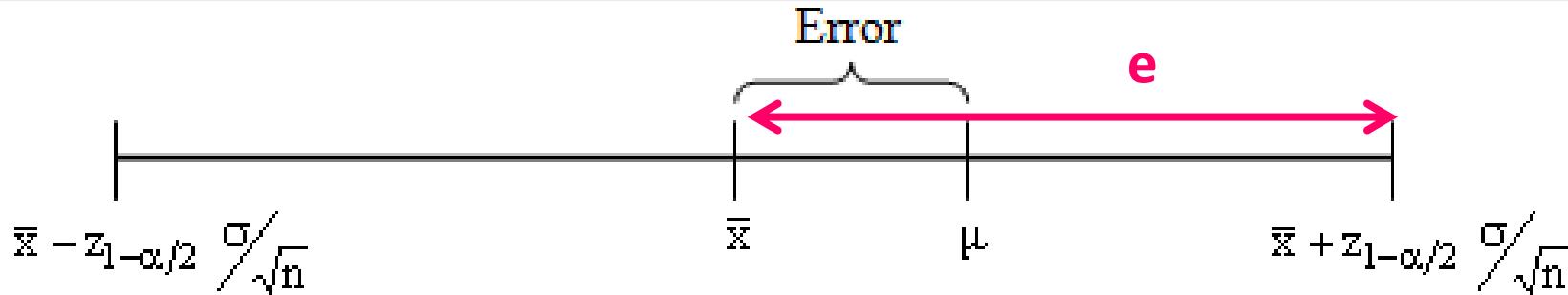
Tamaño de muestra y error máximo para estimar la media de una población normal

El intervalo de confianza $(1-\alpha).100\%$ proporciona grado de precisión de la estimación puntual.

El **tamaño de este error** será el valor absoluto:

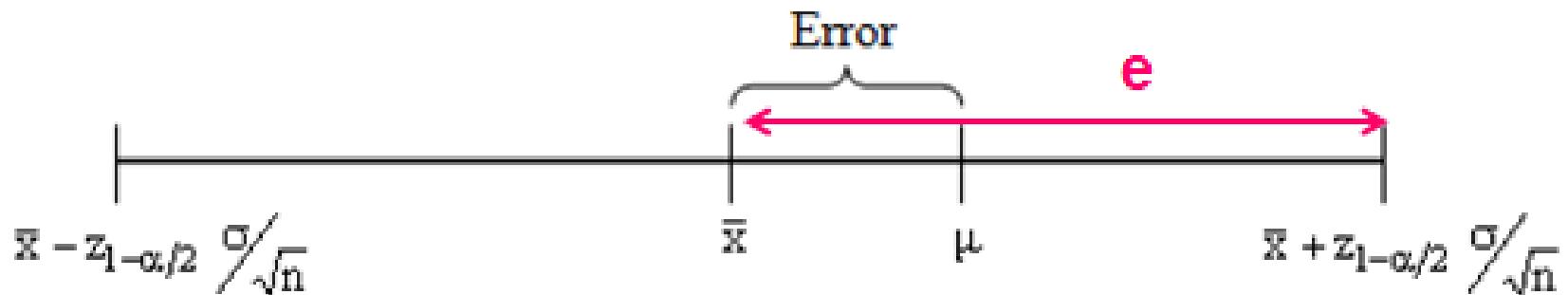
$$|\bar{x} - \mu| = \text{error}$$

Se puede tener una confianza del $(1-\alpha).100\%$ de que **esta diferencia no excederá una cantidad específica e**.



$$|\bar{X} - \mu| = \text{error} < z_{1-\alpha/2} * \frac{\sigma}{\sqrt{n}} = e$$

Error de estimación máximo para estimar la media de una población normal



El **error de estimación máximo** que se cometería al estimar a μ por medio de un IC trabajando con una confianza del $(1 - \alpha)$ 100% y un desvío estándar de σ ,

$$P\left(\left|\bar{X} - \mu\right| = \text{error} < z_{1-\alpha/2} * \frac{\sigma}{\sqrt{n}} = e\right) = 1 - \alpha$$

Tamaño de muestra para estimar la media de una población normal

Para estimar el tamaño de muestra necesario, de manera tal que con una probabilidad $1 - \alpha$, la media muestral, \bar{X} se aleje de la media poblacional μ en a lo sumo e unidades, se recurre a

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1 - \alpha$$

$$P\left(-z_{\alpha/2} * \sigma/\sqrt{n} < \bar{X} - \mu < z_{\alpha/2} * \sigma/\sqrt{n}\right) = 1 - \alpha$$

$$P\left(|\bar{X} - \mu| = error < z_{\alpha/2} * \sigma/\sqrt{n} = e\right) = 1 - \alpha$$

$|\bar{X} - \mu|$ se conoce con el nombre de **error de estimación**, con que se construyó el intervalo de confianza para μ .

Tamaño de muestra para estimar la media de una población normal

A partir de la expresión de **e** es posible obtener el tamaño de muestra **n**.

Entonces, si se utiliza a \bar{X} como estimación puntual de μ se puede tener una confianza del $(1-\alpha)*100\%$ de que el error no excederá una cantidad específica **e** cuando el **tamaño de la muestra** es:

$$n = \left(\frac{z_{1-\alpha/2} \sigma}{e} \right)^2$$

Ejemplo 1

Al gerente de una agencia publicitaria que presta servicios a una compañía radiodifusora le gustaría estimar el tiempo promedio que la audiencia escucha la estación de radio diariamente.

Suponer que el tiempo que un oyente escucha la estación de radio por día es una v.a. normal con un desvío estándar de 45 minutos.

- a) ¿Qué tamaño de muestra se requiere tomar si el gerente de la agencia desea estimar el tiempo promedio que la audiencia escucha la estación por día con una confianza del 95%, cometiendo un error máximo de 5 minutos?
- b) Si no fuese posible tomar un tamaño de muestra como el obtenido en a) y sólo pudiera tomarse las $\frac{3}{4}$ partes del mismo, trabajando con un nivel de confianza de 0.95, ¿qué error máximo se estaría cometiendo?



a)

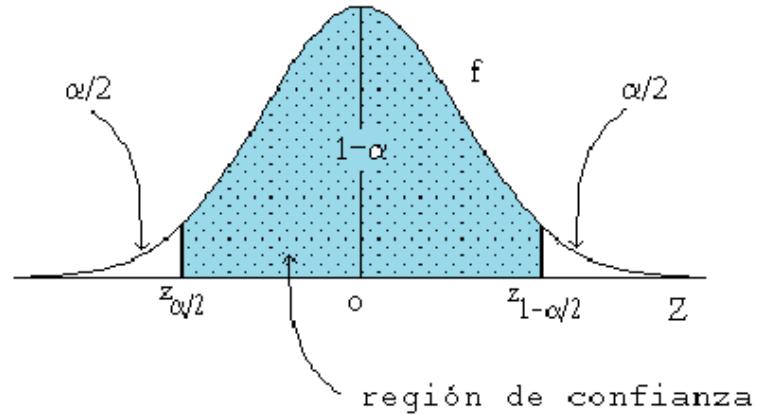
X = “tiempo que un oyente escucha la estación de radio por día (en minutos) ”

$$X \sim N(\mu, \sigma^2 = 45 \text{ minutos}),$$

Datos: $1-\alpha = 0.95$, $e = 5 \text{ minutos}$

$$z_{\alpha/2} = z_{0.025} = -1.96 \quad z_{1-\alpha/2} = z_{0.975} = 1.96$$

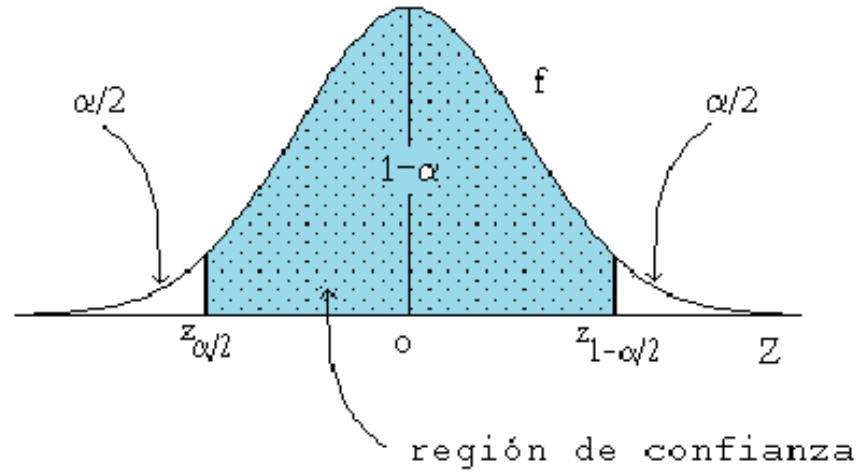
$$n = \left(\frac{z_{1-\alpha/2} \sigma}{e} \right)^2 = \left(\frac{1.96 \sqrt{45}}{5} \right)^2 = 311.17$$



Se requiere seleccionar **aprox. 312 oyentes**, para estimar el tiempo promedio que un oyente escucha la estación de radio por día, trabajando con un **porcentaje de confianza del 95%** y **un error máximo de 5 minutos**.

b)

Datos: $1 - \alpha = 0.95$, $n =$ y $\sigma = 1.5$ minutos



Para estimar el tiempo promedio que un oyente escucha la estación de radio por día, trabajando con un **porcentaje de confianza del 95%** y una muestra de **oyentes**, se estaría cometiendo **un error máximo de 5.77 minutos**.

Ejemplo 2

Como parte de un estudio para estimar salario promedio que percibe un ejecutivo que trabaja en la empresa de servicios informáticos CISET, se eligieron al azar 20 ejecutivos y se construyó un intervalo de confianza. Se sabe que el salario, en pesos, que percibe un ejecutivo de la empresa es una v.a. normal con desvío estándar desconocido. Utilizando la información de la salida del Rcmdr,

```
95 percent confidence interval:  
5622812 7077188
```

```
sample estimates:  
mean of x  
6350000
```

Determinar el **error de estimación** con que se construyó este intervalo.

X = “salario que percibe un ejecutivo que trabaja en la empresa de servicios informáticos CISET” (pesos).

X ~ N(μ, σ²), σ² es desconocido

El intervalo de intervalo de confianza para estimar el salario promedio que percibe un ejecutivo que trabaja en la empresa es:

95 percent confidence
interval:
5622812 7077188

Cuya expresión general es, en este caso: $\left(\bar{X} - t_0 * \frac{S}{\sqrt{n}}, \bar{X} + t_0 * \frac{S}{\sqrt{n}} \right)$

La expresión del error máximo es: $e = t_0 * \frac{S}{\sqrt{n}} = \frac{\text{Amplitud}}{2}$

$$e = \frac{7077188 - 5622812}{2} = 727188$$

El error de estimación con que se construyó es **727188 pesos**

Ejercicio

En una cadena de producción se quiere estimar la longitud media (ℓ) de un cable fabricado mediante un proceso de producción que sigue una distribución normal con una varianza de 0.01cm. Se toman muestras de 16 cables con los siguientes valores de longitud en cm:

4.8, 4.94, 4.75, 4.78, 4.95, 4.91, 4.95, 4.96, 5.02, 4.9, 4.86, 5.01, 5.07, 4.95, 5, 4.84

- a) ¿Cuál será el intervalo de confianza del 95 % para ℓ ?

 - c) ¿Cuál es el error máximo cometido en la estimación de la media poblacional mediante la media muestral?

 - d) ¿Qué deberíamos hacer si quisiéramos reducir este error a 0.03 cm?
-

Intervalo de confianza para la varianza de una población

Sea $X \sim N(\mu, \sigma^2)$, donde σ^2 es el **parámetro desconocido**. Sea X_1, X_2, \dots, X_n una muestra aleatoria de la v.a. X y sea S^2 el **estimador puntual varianza muestral** de σ^2 .

Si la v.a. $X \sim N(\mu, \sigma^2)$, entonces la v.a.

$$\frac{(n-1)S^2}{\sigma^2} = \chi^2_{n-1} = \nu$$

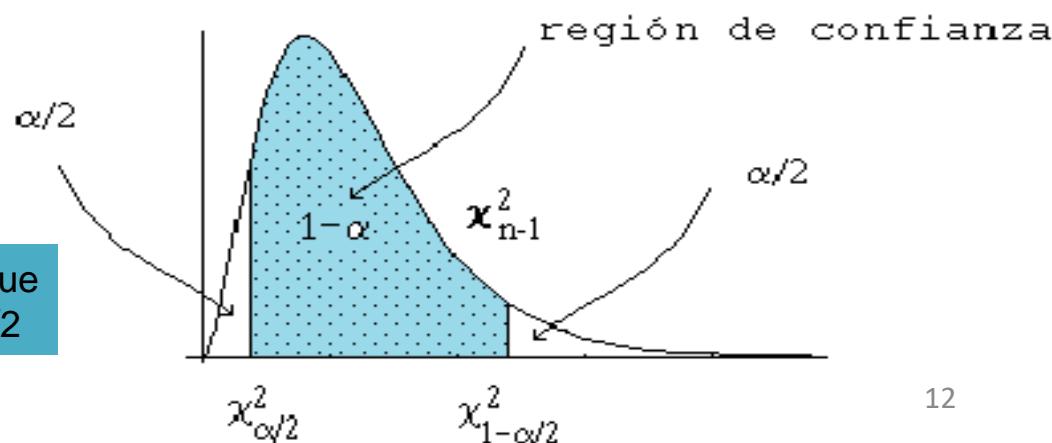
(tiene una distribución ji-cuadrado con $n - 1$ grados de libertad).

Dado $1 - \alpha$, el nivel de confianza elegido para construir el intervalo, es posible encontrar dos valores de la v.a. $\chi^2_{n-1} = \nu$, tales que

$$P(\chi^2_{n-1,\alpha/2} < \chi^2_{n-1} \leq \chi^2_{n-1,1-\alpha/2}) = 1 - \alpha$$

Donde

$\chi^2_{n-1,\alpha/2}$: Valor de la v.a. Ji-cuadrado que deja a su izquierda un área de $\alpha/2$



$\chi^2_{n-1,1-\alpha/2}$: Valor de la v.a. Ji-cuadrado que deja a su izquierda un área de $1 - \alpha/2$

Intervalo de confianza para la varianza de una población

Al sustituir la variable χ^2_{n-1} por la expresión $\frac{(n-1)S^2}{\sigma^2}$ resulta que:

$$P(\chi^2_{\alpha/2, n-1} < \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{1-\alpha/2, n-1}) = 1 - \alpha$$

Al dividir cada término de la desigualdad por $(n-1) S^2$, y a continuación invertir cada término (cambiando por lo tanto el sentido de las desigualdades) se obtiene:

$$P(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2, n-1}} < \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\alpha/2, n-1}}) = 1 - \alpha$$

Si selecciona una muestra aleatoria particular de tamaño n y se calcula la varianza muestral s^2 , el intervalo de confianza del $(1-\alpha)$ 100 % para σ^2 está dado por:

$$\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}}, \frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}} \right) = (\mathbf{a}, \mathbf{b})$$

Intervalo de confianza para el desvío estándar de una población

IMPORTANTE!!!

Un intervalo de confianza del $(1 - \alpha) 100\%$ para el desvío estándar, σ , aplicando la raíz cuadrada a cada punto extremo del intervalo de confianza construido para la varianza, σ^2 :



$$\left(\sqrt{\frac{(n-1)S^2}{\chi_{1-\alpha/2, n-1}^2}}, \sqrt{\frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2}} \right)$$

Intervalo de confianza para la varianza o desvío de una población

Ejemplo 2

Se realizó un estudio respecto de la duración de la licencia anual de vacaciones (en días) que gozan los empleados de una determinada empresa de servicios informáticos. Para ello se eligieron al azar 10 empleados de la empresa y la información recabada se muestra a continuación:

10 11 8 12 13 8 10 12 15 11

Suponer que la duración de la licencia anual de vacaciones de un empleado es una v.a. normalmente distribuida.

Establecer un intervalo de confianza del 99% para estimar el **verdadero desvío estándar** de la duración de la licencia de los empleados de la empresa. Interpretar el intervalo obtenido en términos del problema.

X = “duración de la licencia anual de vacaciones (en días) que goza un empleado (en días) ”

$X \sim N(\mu, \sigma^2)$, donde σ es el parámetro desconocido

Se desea estimar mediante un intervalo de confianza:

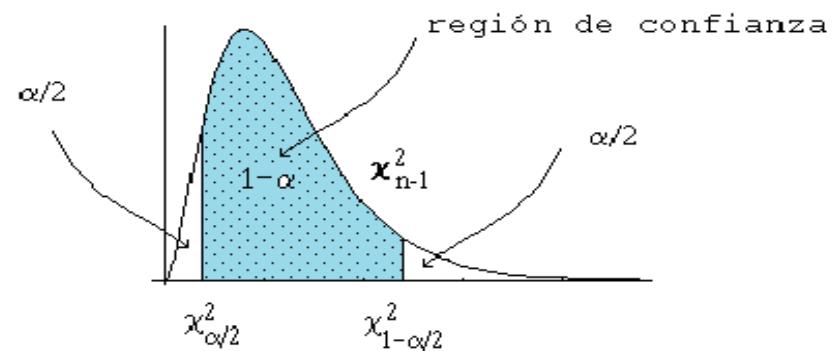
σ = desvío estándar de la duración de la licencia anual de vacaciones (en días) que gozan los empleados ”

Datos: $n = 10, S^2 = 2.16^2 = 4.6656$ días²
 $1 - \alpha = 0.99$ entonces $\alpha = 0.01$ y $\alpha/2 = 0.005$

Intervalo es: $\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}}, \sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}} \right)$

$$\chi^2_{n-1,\alpha/2} = \chi^2_{9,0.005} =$$

$$\chi^2_{n-1,1-\alpha/2} = \chi^2_{9,0.995} =$$



$$\left(\sqrt{\frac{(n-1)S^2}{\chi_{n-1,1-\alpha/2}^2}}, \sqrt{\frac{(n-1)S^2}{\chi_{n-1,\alpha/2}^2}} \right)$$

Reemplazando los datos en:

$$\left(\sqrt{\text{---}}, \sqrt{\text{---}} \right)$$

$$\left(\sqrt{\text{---}}, \sqrt{\text{---}} \right)$$

$$(\text{---}, \text{---}) = (\mathbf{a}, \mathbf{b})$$

El desvío estándar de la duración de la licencia anual de vacaciones (en días) que gozan los empleados --- y --- , con un porcentaje de confianza del --- .

Rstudio: Rcmdr

```
Console Terminal × Background Jobs ×

R - R 4.5.1 - ~/
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribución.

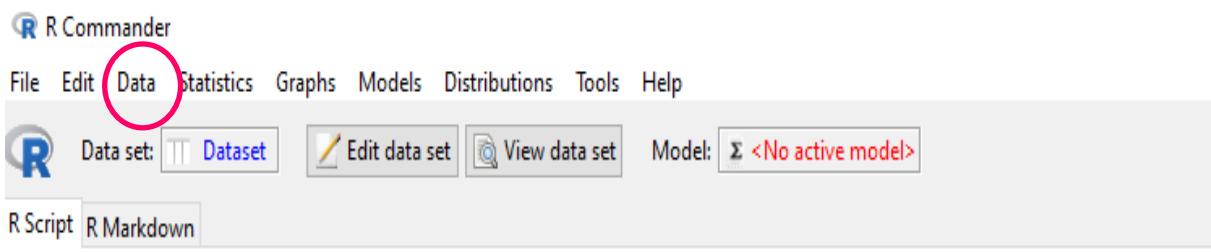
R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line
de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.

[workspace loaded from ~/.RData]

> library(Rcmdr)
```

Cargo la librería “**Rcmdr**”, una vez instalado el package **“Rcmdr”**



```
Dataset <-  
  readXL("C:/Respaldo de archivos/Compact_2020_AG/Compac_2020_agosto/Modelos Estadisticos par  
  rownames=FALSE, header=TRUE, na="", sheet="Hojal", stringsAsFactors=TRUE)  
with(Dataset, t.test(TiempoConexion, alternative='two.sided', mu=0.0,  
  conf.level=.95))  
library(abind, pos=17)  
library(e1071, pos=18)  
numSummary(Dataset[, "TiempoConexion", drop=FALSE], statistics=c("mean",  
  "sd", "quantiles"), quantiles=c(0,.25,.5,.75,1))
```

Clickeo “Data” para importar la base del Excel.

Intervalo de confianza para la varianza o desvío de una población

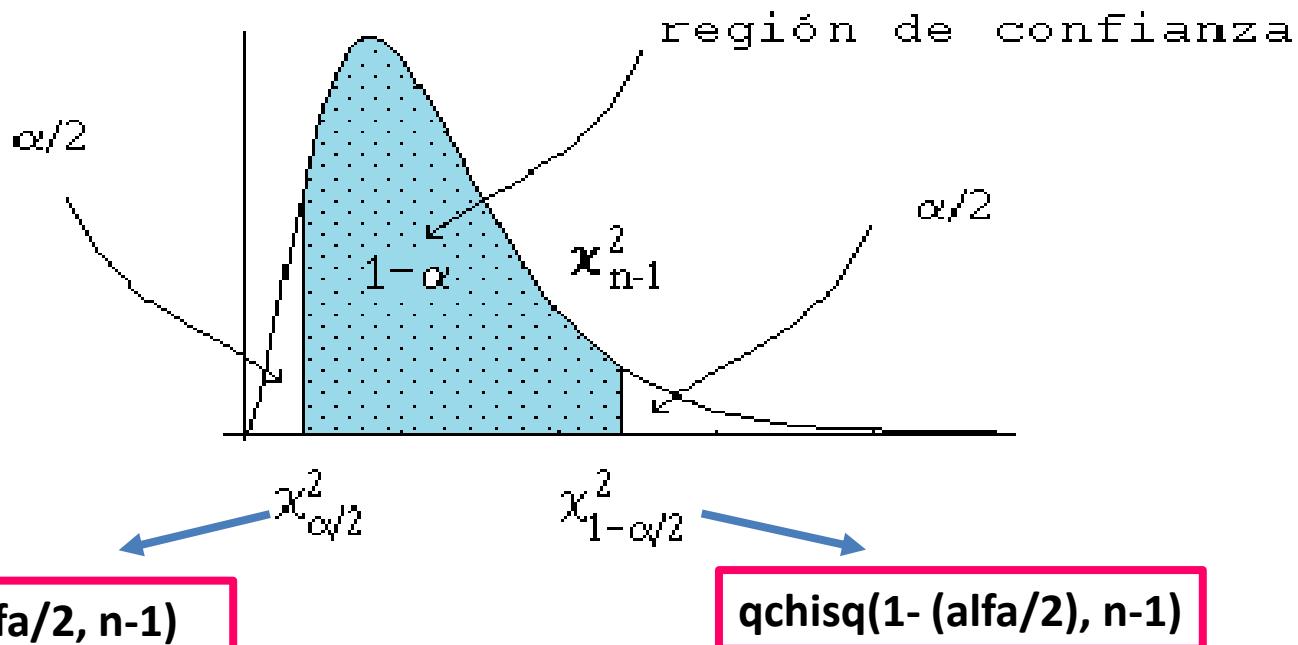
alfa =

n = length(Datos\$variable)

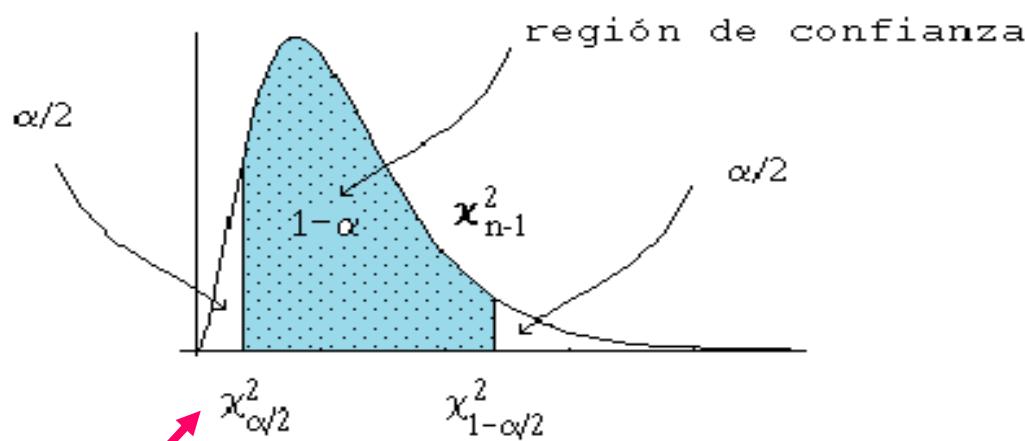
ci =(n-1) *var(Datos\$variable, na.rm =TRUE)/ qchisq(1- (alfa/2), n-1)

cs =(n-1) *var(Datos\$variable, na.rm =TRUE)/ qchisq(alfa/2, n-1)

c(ci,cs)



```
Rcmdr> c(ci,cs)
[1] 1.780464 24.208429
```



```
Rcmdr> qchisq(alfa/2, n-1)
[1] 1.734933
```

```
Rcmdr> qchisq(1-(alfa/2), n-1)
[1] 23.58935
```

Ejercicio

1. Para un estudio de control se desea estimar la alcalinidad media en botellas de agua mineral de una determinada marca . Se puede asumir que la alcalinidad es una v.a. normalmente distribuida con una varianza de 3.84 meq/l^2 . Se eligen al azar 12 botellas, , obteniéndose una alcalinidad promedio de 9 meq/l.
 - a) Estimar puntualmente la alcalinidad media en botellas de agua mineral de dicha marca.
 - b) Construir un intervalo de confianza del 90% para estimar la alcalinidad media en botellas de agua mineral de dicha marca .
 - c) Si se obtiene el mismo valor de alcalinidad promedio pero después de analizar 40 botellas de agua mineral de esa marca, ¿cuáles serán los límites de confianza si se trabaja con un nivel del 90%?
 - d) Comparar los resultados obtenidos en los incisos b) y c).

2. El administrador de una sucursal de un banco desea estimar el promedio y la variabilidad de la cantidad de dinero existente en las cuentas de ahorro de los clientes del banco. Para ello se seleccionó una muestra aleatoria de 25 depositantes que arrojó un monto promedio de 4750\$ y un desvío de 1200\$. Se puede asumir que la cantidad de dinero existente en una cuenta de ahorro de un cliente del banco es una v.a. normalmente distribuida.

- a) Estimar puntualmente la variabilidad de la cantidad de dinero existente en una cuenta de ahorro de un cliente del banco.
- b) Construir un intervalo de confianza del 95% para estimar la cantidad promedio de dinero existente en las cuentas de ahorro de los clientes del banco.
- c) Construir un intervalo de confianza del 95% para estimar el desvío de la cantidad de dinero existente en las cuentas de ahorro de los clientes del banco.