Universidad del País Vasco
Euskal Herriko Unibertsitatea
BILBOKO INGENIARITZA ESKOLA
ESCUELA DE INGENIERÍA DE BILBAO

MÉTODOS ESTADÍSTICOS DE LA INGENIERÍA

QUINTA PRÁCTICA DE ORDENADOR

ESTIMACIÓN

1 DISTRIBUCIONES

Prefijos

Función de probabilidad/densidad	d
Función de distribución	p
Crear valores aleatorios	r
Función cuantil	q

Distribuciones continuas

Normal	norm
χ^2	chisq
t de Student	t
F de Snedecor	f

Universidad del País Vasco Euskal Herriko Unibertsitatea

BILBOKO INGENIARITZA ESKOLA
ESCUELA DE INGENIERÍA DE BILBAO

1 TABLA RESUMEN DE LAS DISTRIBUCIONES CONTINUAS

Muestreo/Estimación				
Distribución	**Función de densidad** $f(x)$**:** $P(a<X<b)=\int_a^b f(x)dx$ $\forall a,b\in\mathbb{R} \quad \vee \quad a=-\infty, b=\infty$	**Función de distribución** $F(x)$**:** $F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)dt \quad \forall x\in\mathbb{R}$	**Cuantiles**	**Muestras aleatorias**
Normal: $X\sim N(\mu,\sigma)$	`dnorm(x,μ,σ)`	`pnorm(x,μ,σ)`	`qnorm(pr,μ,σ)`	`rnorm(o,μ,σ)`
χ2 (Pearson): $X\sim\chi^2_\nu$	`dchisq(x,ν)`	`pchisq(x,ν)`	`qchisq(pr,ν)`	`rchisq(o,ν)`
t de Student: $X\sim t_\nu$	`dt(x,ν)`	`pt(x,ν)`	`qt(pr,ν)`	`rt(o,ν)`
F de Snedecor: $X\sim F_{n,m}$	`df(x,n,m)`	`pf(x,n,m)`	`qf(pr,n,m)`	`rf(o,n,m)`

1) Notación: pr: vector de probabilidad; o: número de datos; ν: grados de libertad.

2) En las funciones `p` y `q` se puede añadir el argumento `lower.tail=F`, por defecto R tiene definido `lower.tail=T`, añadiendo el argumento `lower.tail=F`, se calcula la probabilidad: $1-F(x)=P(X>x) \quad \forall x$

Universidad del País Vasco Euskal Herriko Unibertsitatea

BILBOKO INGENIARITZA ESKOLA ESCUELA DE INGENIERÍA DE BILBAO

2 FUNCIONES PARA LA ESTIMACIÓN (COMANDOS)

INTERVALO DE CONFIANZA	COMANDO
Media (μ)	t.test(muestra)$conf
Varianza (σ^2)	-----
División entre varianzas(σ_1^2/σ_2^2)	var.test(muestra 1, muestra 2)$conf
Diferencia de medias (μ_1-μ_2)	t.test(muestra 1, muestra 2)$conf
Proporción (p)	prop.test(x,n)$conf

OPCIONES	COMANDO
Nivel de confianza	conf.level
Igualdad de varianzas ($\sigma_1^2 = \sigma_2^2$) (Para tener en cuenta que $\sigma_1^2/\sigma_2^2 = 1$)	var.equal=T (A emplear en la diferencia de medias)

Universidad del País Vasco Euskal Herriko Unibertsitatea

BILBOKO INGENIARITZA ESKOLA ESCUELA DE INGENIERÍA DE BILBAO

3

EJEMPLOS DE ESTIMACIÓN

1. **En una planta de potabilización se han medido durante 10 días los niveles de cloro en las aguas. Los resultados obtenidos son: 2.2, 1.9, 1.7, 1.6, 1.7, 1.8, 1.7, 1.9, 2.0, 2.0. La distribución del nivel de cloro se considera normal.**

a) X = "Nivel de cloro". Calcula la estimación puntual de μ y σ del nivel de cloro.

```
>nivelcloro <- c(2.2,1.9,1.7,1.6,1.7,1.8,1.7,1.9,2.0,2.0)
>nivelcloro
[1] 2.2 1.9 1.7 1.6 1.7 1.8 1.7 1.9 2.0 2.0

>mean(nivelcloro)#Estimación puntual de la media
[1] 1.85

>var(nivelcloro)#Estimación puntual de la varianza
[1] 0.03388889

>sd(nivelcloro)#Estimación puntual de la desviación típica
[1] 0.1840894
```

3 EJEMPLOS DE ESTIMACIÓN

1. En una planta de potabilización se han medido durante 10 días los niveles de cloro en las aguas. Los resultados obtenidos son: 2.2, 1.9, 1.7, 1.6, 1.7, 1.8, 1.7, 1.9, 2.0, 2.0. La distribución del nivel de cloro se considera normal.

b) El intervalo de confianza de la media del nivel de cloro, al nivel de confianza del 95%.

```
>t.test(nivelcloro)$conf #σ desconocida distribución normal  (R emplea la distribución t de Student)
[1] 1.71831 1.98169
attr(,"conf.level")
[1] 0.95

#Escribiendo la expresión en R
>x <- mean(nivelcloro)
>S <- sd(nivelcloro)
>0.05/2
[1] 0.025

> IC95 <-c(x-qt(0.025,9,lower.tail=F)*S/sqrt(10), x+qt(0.025,9,lower.tail=F)*S/sqrt(10))
> IC95
[1] 1.71831 1.98169
```

$$I_{\mu}^{1-\alpha} = \left[\bar{x} - t_{n-1;\alpha/2}.\, {}^{S}/_{\sqrt{n}} , \bar{x} + t_{n-1;\alpha/2}.\, {}^{S}/_{\sqrt{n}}\right]$$

3 EJEMPLOS DE ESTIMACIÓN

1. **En una planta de potabilización se han medido durante 10 días los niveles de cloro en las aguas. Los resultados obtenidos son: 2.2, 1.9, 1.7, 1.6, 1.7, 1.8, 1.7, 1.9, 2.0, 2.0. La distribución del nivel de cloro se considera normal.**

c) El intervalo de confianza de la media del nivel de cloro, al nivel de confianza del 99%.

```
>t.test(nivelcloro, conf.level=0.99)$conf #σ desconocida distribución normal (R emplea la t de Student)
[1] 1.660814 2.039186
attr(,"conf.level")
[1] 0.99

#Escribiendo la expresión en R
>x <- mean(nivelcloro)
>S <- sd(nivelcloro)
>0.01/2
[1] 0.005
```

$$I_{\mu}^{1-\alpha}=\left[\bar{x}-t_{n-1;\alpha/2}.\,S/\sqrt{n}\,,\,\bar{x}+t_{n-1;\alpha/2}.\,S/\sqrt{n}\right]$$

```
> IC99 <-c(x-qt(0.005,9,lower.tail=F)*S/sqrt(10), x+qt(0.005,9,lower.tail=F)*S/sqrt(10))
> IC99
[1] 1.660814 2.039186
```

3

EJEMPLOS DE ESTIMACIÓN

1. **En una planta de potabilización se han medido durante 10 días los niveles de cloro en las aguas. Los resultados obtenidos son: 2.2, 1.9, 1.7, 1.6, 1.7, 1.8, 1.7, 1.9, 2.0, 2.0. La distribución del nivel de cloro se considera normal.**

d) **El intervalo de confianza de la varianza del nivel de cloro con un nivel de confianza del 95%.**

No hay comando. Se ha de escribir la expresión en R

μ **desconocida, distribución normal**

$$I_{\sigma^2}^{1-\alpha} = \left[\frac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}\right]$$

```
#Escribiendo la expresión en R
>ICsigma95<-c(9*var(nivelcloro)/qchisq(0.975,9),9*var(nivelcloro)/qchisq(0.025,9))
>ICsigma95
[1] 0.01603342 0.11294667
```

e) **El intervalo de confianza de la varianza con un nivel de confianza del %99**

```
>ICsigma99<-c(9*var(nivelcloro)/qchisq(0.995,9),9*var(nivelcloro)/qchisq(0.005,9))
>ICsigma99
[1] 0.01292956 0.17579931
```

3

EJEMPLOS DE ESTIMACIÓN

2. Realizar la estimación puntual y estimar mediante un intervalo de confianza a un 95% de nivel de confianza, la proporción de piezas defectuosas de un lote que sigue una distribución binomial. Para ello se escogen al azar 150 piezas, observándose que 12 de ellas son defectuosas.

X: “Número de piezas defectuosas”

Estimación puntual

>12/150#estimación puntual de p
[1] 0.08

Intervalo de confianza

>prop.test(12,150)$conf #R emplea la distribución χ^2 (en nuestro caso emplearemos la distribución nomal)
[1] 0.04388586 0.13863413
attr(,"conf.level")
[1] 0.95

$$I_p^{1-\alpha} = \left[\hat{p} - z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}\right]$$

#Escribiendo la expresión en R
>p<-0.08;q<-0.92
>P95 <-c(p-qnorm(0.025,0,1,lower.tail=F)*sqrt(p*q/150), p+qnorm(0.025,0,1,lower.tail=F)*sqrt(p*q/150))
>P95
[1] 0.03658484 0.12341516

3 EJEMPLOS DE ESTIMACIÓN

3. Se quiere saber si existen diferencias significativas en la facturación de dos tiendas de joyería de una misma cadena. Para ello se eligieron al azar 11 días en los que se contabilizaron las ventas en la joyería A y otros 10 días en la joyería B. Los datos se encuentran en el archivo Ventas.txt. X: "Ventas de la joyería A" , Y: "Ventas de la joyería B"

a) Estimación puntual de la media y la varianza de las ventas en las joyerías A y B.

```
>datos <- read.table("Ventas.txt",header=T)#Lectura de datos
>attach(datos)#Datos disponibles
>VentasA
[1] 1320 1495 990 1250 12900 1900 1500 1100 1250 1100 1930
>VentasB
[1] 1110 1405 985 1290 1300 1705 1200 1105 1150 1210 NA

>mean(VentasA)
[1] 2430.455
>mean(VentasB,na.rm=T) #Para no tener en cuenta los datos perdidos (NA)
[1] 1246
>var(VentasA)
[1] 12151922
> var(VentasB,na.rm=T) #Para no tener en cuenta los datos perdidos (NA)
[1] 39993.33
```

3 EJEMPLOS DE ESTIMACIÓN

3. **Se quiere saber si existen diferencias significativas en la facturación de dos tiendas de joyería de una misma cadena. Para ello se eligieron al azar 11 días en los que se contabilizaron las ventas en la joyería A y otros 10 días en la joyería B. Los datos se encuentran en el archivo Ventas.txt. X: "Ventas de la joyería A" , Y: "Ventas de la joyería B"**

b) **El intervalo de confianza de la división de las varianzas de las ventas de las joyerías con un nivel de confianza del 95%**

>var.test(VentasA,VentasB)$conf

[1] 76.65465 1148.23288

attr(,"conf.level")

[1] 0.95#Se puede concluir que las varianzas son distintas , puesto que el valor de 1 no se halla dentro del intervalo

c) **El intervalo de confianza de la diferencia de medias de las ventas de las joyerías con un nivel de confianza del 95%**

>t.test(VentasA,VentasB,var.equal=F)$conf
[1] -1159.397 3528.307
attr(,"conf.level")
[1] 0.95

#Como se ha visto que las varianzas son distintas, se elige var.equal=F , aunque no es necesario escribirlo puesto que en R se encuentra por defecto esta opción. Por tanto, sólo habría que escribirlo en el caso de var.equal=T.

Universidad del País Vasco Euskal Herriko Unibertsitatea
BILBOKO INGENIARITZA ESKOLA
ESCUELA DE INGENIERÍA DE BILBAO

3 EJEMPLOS DE ESTIMACIÓN

4. Genere una muestra aleatoria de tamaño 100 de una variable aleatoria N(10,2).

```
>valores<-rnorm(100,10,2) #Creación de la muestra
```

a) Calcule el intervalo de confianza de la media con un nivel de confianza del 95%.

```
>t.test(valores)$conf
[1] 9.786608 10.508501
attr(,"conf.level")
[1] 0.95
```

b) ¿Cuántos valores se quedan fuera del intervalo?

```
>N1<-length(which(valores<9.786608))
>N2<-length(which(valores>10.508501))
>N1+N2
[1] 78
```