# Estadística Bayesiana: Clase 6

Juan Carlos Correa

16 de marzo de 2021

# Utilidad de las distribuciones apriori no informativas en la estadística clásica

## Intervalo de confianza para la proporción vía TCL

Este es el intervalo propuesto en la mayoría de textos básicos en estadística (Canavos, 1988; Wonnacott y Wonnacott, 1979; Roussas, 1973; Walpole, 1992; Meyer, 1986; Mood et al., 1974)

$$\left(\hat{\pi} - z_{\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + z_{\alpha/2}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right)$$

# Intervalo clásico vía estadística bayesiana

## Intervalo de confianza para la proporción vía apriori de Laplace

El intervalo es $(LI, LS)$ donde los límites se obtienen como soluciones de las siguientes ecuaciones

$$\int_0^{LI} \xi\left(\pi \,|Datos\right) = \frac{\alpha}{2}$$

$$\int_0^{LS} \xi\left(\pi \,|Datos\right) = 1 - \frac{\alpha}{2}$$

donde $\xi\left(\pi \,|Datos\right)$ es una $Beta\left(\sum_i x_i + 1;\; n - \sum_i x_i + 1\right)$.

## Intervalo de confianza para la proporción vía apriori de Jeffreys

El intervalo es $(LI, LS)$ donde los límites se obtienen como soluciones de las siguientes ecuaciones

$$\int_0^{LI} \xi\left(\pi\,|Datos\right) = \frac{\alpha}{2}$$

$$\int_0^{LS} \xi\left(\pi\,|Datos\right) = 1 - \frac{\alpha}{2}$$

donde $\xi\left(\pi\,|Datos\right)$ es una $Beta\left(\sum_i x_i + \frac{1}{2};\; n - \sum_i x_i + \frac{1}{2}\right)$.

# Comparación de los tres intervalos

```
# Intervalo de confianza para la proporción

intervalo.vía.tcl<-function(x,n,nivel=0.95){
# x: Número de éxitos
# n: tamaño muestral
# nivel: nivel de confianza pre-establecido en el 95 porciento

if(x==0) x<-0.5
if(x==n) x<-n-0.5

pi.gorro<-x/n
error<-sqrt(pi.gorro*(1-pi.gorro)/n)

percentil<-qnorm(nivel+(1-nivel)/2)

LI<-pi.gorro-percentil*error
if(LI<0) LI<-0

LS<-pi.gorro+percentil*error
if(LS>1) LS<-1

list(LI=LI,LS=LS)
} #Fin intervalo.vía.tcl
```

```
intervalo.vía.Laplace<-function(x,n,nivel=0.95){
# x: Número de éxitos
# n: tamaño muestral
# nivel: nivel de confianza pre-establecido en el 95 porciento

percentil1<-(1-nivel)/2
percentil2<-nivel+(1-nivel)/2

LI<-qbeta(percentil1,x+1,n-x+1)
LS<-qbeta(percentil2,x+1,n-x+1)

list(LI=LI,LS=LS)
} #Fin intervalo.vía.Laplace
##############################################################
intervalo.vía.Jeffreys<-function(x,n,nivel=0.95){
# x: Número de éxitos
# n: tamaño muestral
# nivel: nivel de confianza pre-establecido en el 95 porciento

percentil1<-(1-nivel)/2
percentil2<-nivel+(1-nivel)/2

LI<-qbeta(percentil1,x+1/2,n-x+1/2)
LS<-qbeta(percentil2,x+1/2,n-x+1/2)

list(LI=LI,LS=LS)
} #Fin intervalo.vía.Jeffreys
```

```
# Comparación de los tres métodos
# Asumamos que la verdadera proporción es 0.3

n<-15
x<-0:15
cae<-function(inter,p.ver)if(p.ver>inter[1] & p.ver<inter[2])
res<-1 else res<-0

probas.verdaderas<-dbinom(x,n,prob=0.3)
intervalos.resu<-matrix(unlist(sapply(x,intervalo.vía.tcl,n)),
ncol=2,byrow=T)
amplitud.intervalo<-intervalos.resu[,2]-intervalos.resu[,1]
longi.media1<-sum(amplitud.intervalo*probas.verdaderas)

tempo1<-apply(intervalos.resu,1,cae,0.3)
nivel.real1<-sum(tempo*probas.verdaderas)
```

```
intervalos.resu<-matrix(unlist(sapply(x,intervalo.vía.Laplace,n)),
ncol=2,byrow=T)
amplitud.intervalo<-intervalos.resu[,2]-intervalos.resu[,1]
longi.media2<-sum(amplitud.intervalo*probas.verdaderas)
tempo<-apply(intervalos.resu,1,cae,0.3)
nivel.real2<-sum(tempo*probas.verdaderas)

intervalos.resu<-matrix(unlist(sapply(x,intervalo.vía.Jeffreys,n)),
ncol=2,byrow=T)
amplitud.intervalo<-intervalos.resu[,2]-intervalos.resu[,1]
longi.media3<-sum(amplitud.intervalo*probas.verdaderas)
tempo<-apply(intervalos.resu,1,cae,0.3)
nivel.real3<-sum(tempo*probas.verdaderas)
```

```
> longi.media1
[1] 0.4380546
> nivel.real1
[1] 0.9647319
>
> longi.media2
[1] 0.4118484
> nivel.real2
[1] 0.9800099
>
> longi.media3
[1] 0.4166257
> nivel.real3
[1] 0.9494899
```

| Intervalo | Longitud promedio | Nivel de confianza real |
|---|---|---|
| TCL | 0.4380546 | 0.9647319 |
| Laplace | 0.4118484 | 0.9800099 |
| Jeffreys | 0.4166257 | 0.9494899 |

## Caso Poisson

Si $X(\lambda)$ su función de probabilidad es

$$f\left(x\left|\lambda\right.\right) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$f\left(x\left|\lambda\right.\right) \propto \lambda^x e^{-\lambda}$$

para $\lambda > 0$ y $x = 0, 1, 2, \ldots$

## Distribución apriori no informativa Laplace para el parámetro $\lambda$ de la Poisson

$$\xi(\lambda) = Uniforme\left(0, \infty\right)$$

$$\xi(\lambda) \propto 1$$

Es impropia!!

## Si tenemos una muestra de tamaño $n$ de esta Poisson

Si tenemos la muestra aleatoria $X_1, \ldots X_n$

- La verosimilitud

$$L\left(\lambda \left| Datos \right.\right) \propto \lambda^{\sum_i x_i} \exp\left(-n\lambda\right)$$

- La apriori no informativa de Laplace

$$\xi(\lambda) \propto 1$$

- La aposteriori

$$\xi(\lambda \left| Datos \right.) \propto \xi(\lambda) L\left(\lambda \left| Datos \right.\right) \propto \lambda^{\sum_i x_i + 1 - 1} \exp\left(-n\lambda\right)$$

Este es el kernel de una distribución $Gamma\left(\sum_i x_i + 1, n\right)$

## Distribución apriori no informativa Jeffreys para el parámetro $\lambda$ de la Poisson

$$
\begin{aligned}
f(x|\lambda) &= \frac{\lambda^x e^{-\lambda}}{x!} \\
\log(f(x|\lambda)) &= x\log(\lambda) - \lambda - \log(x!) \\
\frac{d\,\log(f(x|\lambda))}{d\,\lambda} &= \frac{x}{\lambda} - 1 \\
\frac{d^2\,\log(f(x|\lambda))}{d\,\lambda} &= -\frac{x}{\lambda^2} \\
\mathcal{I} = -\mathcal{E}\left[\frac{\lceil^{\in}\,\log(\{(\S|\lambda))}{\lceil\,\lambda}\right] &= \frac{E(x)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}
\end{aligned}
$$

Por lo tanto

$$\xi(\lambda) \propto \lambda^{-1/2}$$

## Si tenemos una muestra de tamaño $n$ de esta Poisson en el caso de la apriori no informativa de Jeffreys para $\lambda$

Si tenemos la muestra aleatoria $X_1, \ldots X_n$

- La verosimilitud

$$L\left(\lambda \left| Datos\right.\right) \propto \lambda^{\sum_i x_i} \exp\left(-n\lambda\right)$$

- La apriori no informativa de Jeffreys

$$\xi(\lambda) \propto \lambda^{-1/2}$$

- La aposteriori

$$\xi(\lambda \left| Datos\right.) \propto \xi(\lambda) L\left(\lambda \left| Datos\right.\right) \propto \lambda^{\sum_i x_i + 1 - 1 - 1/2} \exp\left(-n\lambda\right)$$

Este es el kernel de una distribución $Gamma\left(\sum_i x_i + 1/2, n\right)$

## Intervalos de confianza para $\lambda$ (la media de la Poisson)

## Método basado en el Teorema Central del Límite (T.C.L.)

Si el tamaño muestral es lo suficientemente grande, podemos aplicar el teorema central del límite.

$$\left(\bar{X} - z_{\alpha/2}\frac{s}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

donde $s^2 = 1/n \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$. Este es el intervalo propuesto en la mayoría de textos básicos en estadística (Canavos, 1988; Wonnacott y Wonnacott, 1979; Roussas, 1973; Walpole, 1992; Meyer, 1986; Mood et al., 1974)

## Método basado en la Máxima Verosimilitud

Se sabe que si $\hat{\theta}$ es el estimador máximo verosímil para $\theta$ (puede ser un vector), bajo ciertas condiciones suaves (Serfling, 1980), entonces $\hat{\theta} \sim (\theta, I^{-1}(\theta))$, con $I(\theta)$ siendo la matriz de información de Fisher. Entonces, en el caso Poisson

$$\left(\bar{X} - z_{\alpha/2}\frac{\sqrt{\bar{X}}}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sqrt{\bar{X}}}{\sqrt{n}}\right)$$

Hay muchos más intervalos de confianza en el caso Poisson!

## Intervalos de Confianza Cuando no hay Eventos

Cuando las tasas de ocurrencia son muy bajas es común tener muestras donde todas las observaciones son ceros. El problema entonces es construir el intervalo de confianza para $\lambda$. Asumamos que $x_1 = 0, x_2 = 0, \cdots, x_n = 0$, con $n$ fijo es el resultado de la muestra aleatoria. Podemos entonces proceder así:

$$P(X_1 = 0\ , X_2, \cdots, X_n = 0) = e^{-n\lambda}$$

Entonces el límite superior, $LS$, será calculado como

$$\max_{\lambda}\ e^{-n\lambda} \geq \alpha$$

lo cual produce $LS = -\frac{\ln(\alpha)}{n}$. El intervalo será $(0, LS)$.

## Método basado en la apriori no informativa de Laplace

$$(LI, LS)$$

donde $LI$ se obtiene como el cuantil $\alpha/2$ de una distribución $Gamma\,(\sum_i x_i + 1; n)$ y $LS$ se obtiene como el cuantil $1-\alpha/2$ de una distribución $Gamma\,(\sum_i x_i + 1; n)$.

## Método basado en la apriori no informativa de Jeffreys

$$(LI, LS)$$

donde $LI$ se obtiene como el cuantil $\alpha/2$ de una distribución $Gamma\left(\sum_i x_i + 1; n\right)$ y $LS$ se obtiene como el cuantil $1-\alpha/2$ de una distribución $Gamma\left(\sum_i x_i + 1/2; n\right)$.

# Comparación vía simulación de los cuatro métodos para hallar intervalos de confianza para $\lambda$

```
# Intervalos de confianza para la lambda

intervalo.vía.tcl<-function(x,nivel=0.95){
# x: vector con los eventos
# nivel: nivel de confianza pre-establecido en el 95 porciento

n<-length(x)
y<-sum(x)
if(y==0){
 LS<--log(1-nivel)/n
LI<-0
}
else{
media<-mean(x)
desvi<-sd(x)
Z<-qnorm(1-(1-nivel)/2)
LI<-media-Z*desvi/sqrt(n)
if(LI<0)LI<-0
LS<-media+Z*desvi/sqrt(n)
}

list(LI=LI,LS=LS)
} #Fin intervalo.vía.tcl
```

```
intervalo.vía.mv<-function(x,nivel=0.95){
# x: vector con los eventos
# nivel: nivel de confianza pre-establecido en el 95 porciento

n<-length(x)
y<-sum(x)
if(y==0){
 LS<--log(1-nivel)/n
LI<-0
}
else{
media<-mean(x)
Z<-qnorm(1-(1-nivel)/2)
LI<-media-Z*sqrt(media/n)
if(LI<0)LI<-0

LS<-media+Z*sqrt(media/n)
}

list(LI=LI,LS=LS)
} #Fin intervalo.vía.mv
```

```
intervalo.vía.Laplace<-function(x,nivel=0.95){
# x: vector con los eventos
# nivel: nivel de confianza pre-establecido en el 95 porciento

percentil1<-(1-nivel)/2
percentil2<-nivel+(1-nivel)/2
y<-sum(x)

LI<-qgamma(percentil1,y+1,rate=n)

LS<-qgamma(percentil2,y+1,rate=n)

list(LI=LI,LS=LS)
} #Fin intervalo.vía.Laplace
```

```
intervalo.vía.Jeffreys<-function(x,nivel=0.95){
# x: Número de éxitos
# nivel: nivel de confianza pre-establecido en el 95 porciento

percentil1<-(1-nivel)/2
percentil2<-nivel+(1-nivel)/2

y<-sum(x)
LI<-qgamma(percentil1,y+1/2,rate=n)

LS<-qgamma(percentil2,y+1/2,rate=n)

list(LI=LI,LS=LS)
} #Fin intervalo.vía.Jeffreys
```

```
# Comparación de los cuatro métodos
# Asumamos que la verdadera lambda es 2

lambda.verd<-2
n<-15
Nsim<-1000

muestras<-matrix(rpois(n*Nsim,2),ncol=n)

res1<-matrix(unlist(apply(muestras,1,intervalo.vía.tcl)),
ncol=2,byrow=T)
res2<-matrix(unlist(apply(muestras,1,intervalo.vía.mv)),
ncol=2,byrow=T)
res3<-matrix(unlist(apply(muestras,1,intervalo.vía.Laplace)),
ncol=2,byrow=T)
res4<-matrix(unlist(apply(muestras,1,intervalo.vía.Jeffreys)),
ncol=2,byrow=T)
```

```
cae<-function(inter,lambda.ver)if(lambda.ver>inter[1] &
lambda.ver<inter[2]) res<-1 else res<-0

longi.media1<-mean(res1[,2]-res1[,1])
nivel.real1<-mean(apply(res1,1,cae,lambda.verd))

longi.media2<-mean(res2[,2]-res2[,1])
nivel.real2<-mean(apply(res2,1,cae,lambda.verd))

longi.media3<-mean(res3[,2]-res3[,1])
nivel.real3<-mean(apply(res3,1,cae,lambda.verd))

longi.media4<-mean(res4[,2]-res4[,1])
nivel.real4<-mean(apply(res4,1,cae,lambda.verd))
```

```
> longi.media1
[1] 1.415656
> nivel.real1
[1] 0.931
>
> longi.media2
[1] 1.434487
> nivel.real2
[1] 0.936
>
> longi.media3
[1] 1.45418
> nivel.real3
[1] 0.953
>
> longi.media4
[1] 1.442293
> nivel.real4
[1] 0.963
```