Estadística Bayesiana: Clase 9

Juan Carlos Correa

25 de marzo de 2021

Indice de Dispersión de Fisher

Esta es la más antigua prueba análitica diseñada específicamente para la Poisson (Santner y Duffy, pp.93). El estadístico de prueba es

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(Y_i - \bar{Y}\right)^2}{\bar{Y}}$$

Este estadístico se distribuye asintóticamente χ^2 con $n-1$ grados de libertad.

Podemos modelar los goles del equipo local con una distribución Poisson?

```
goles.local.2001<-scan()
0 0 0 1 2 2 2 0 1 0 3 2 1 0 1 5 1 3 2 2 0 0 0 2 1 3 3 1 0 2 2 2
0 1 0 1 1 1  4 0 3 2 1 1 1 2 1 3 1 3 1 3 2 1 2 1 5 0 2 3 1 1 2
2 1 0 3 2 2 2 6 3 2 2 3 2 2 3 2 2 2 3 2 1 2 2 1 1 0 1 1 2 0 1 3
1 3 0 2 2 4 1 2 1 3 3 0 0 2 2 0 3 2 2 1 2 0 3 2 1 0 0  5 2 1 0 2
3 3 1 2 1 2 2 0 2 1 3 1 0 1 0 1 0 1 1 1 0 2 2 3  2 4 2 1 0 2 2 1
2 0 1 1 2 2 4 0 2 1 2 3 1 1 1 4 2 3 0 2 4 0 1 1 0 2 1 2 0 0 2 3 4
3 2 1 2 3 1 0 1 2 1 1 0 2 0 1 1 3 0 4 1 0 0 3 0 0 0 3 2 0 2 1 2
3 2 1 0 0 2 2 1 1 2 2 1 1 1 1 1 2 1 2 4 2 1 0 5 1 1 0

 m<-mean(goles.local.2001)
v<-var(goles.local.2001)
n<-length(goles.local.2001)
Fisher<-(n-1)*v/m
pchisq(Fisher,n-1,lower.tail=F)
[1] 0.8392992
> m
[1] 1.553785
> v
[1] 1.416096
> n
[1] 251
```

Ejemplo: Poisson

Sea $y_1, \cdots y_n$ una muestra aleatoria de una $Poisson(\lambda)$. Supongamos también que la apriori es una $Gamma(1,1)$. Por lo tanto la aposterior será $Gamma(1+\sum_{i=1}^{n} y_i, n+1)$.
El estimador bayesiano para λ

- bajo la función de pérdida cuadrática es

$$\hat{\lambda} = \frac{1+\sum_{i=1}^{n} y_i}{n+1}$$

- bajo la función de pérdida escalonada

$$\hat{\lambda} = \frac{\alpha^* - 1}{\beta^*} = \frac{\sum_{i=1}^{n} y_i}{n+1} \text{ si } \alpha^* \geq 1$$

La siguiente función en R calcula los tres estimadores, bajo el supuesto de una aprori $Gamma(\alpha_0, \beta_0)$:

```
calcula.estimadores.poisson<-function(alfa0,beta0,x,n=lenght(x))
{
alfa1<-alfa0+sum(x)
beta1<-beta0+n
estimador.fpc<-alfa1/beta1
estimador.fpa<-qgamma(0.5,alfa1,beta1)
estimador.fpe<-(alfa1-1)/beta1
list(estimador.fpc=estimador.fpc,
      estimador.fpa=estimador.fpa,
      estimador.fpe=estimador.fpe)
}
```

La utilización será

```
> calcula.estimadores.poisson(1,1,16,n=4)

$estimador.fpc
[1] 3.4

$estimador.fpa
[1] 3.333571

$estimador.fpe
[1] 3.2
```

Estimador generalizado de máxima verosimilitud

Si tal valor de w existe para todo valor de x,

$$\xi\left[\hat{w}(x)|x\right] = \sup_{w \in \Omega} \xi(w|x)$$

entonces decimos que el estimador $\hat{w}(X)$ es un estimador generalizado de máxima verosimilitud de W.

Regiones de Credibilidad

Los intervalos de confianza clásicos frecuentemente son malinterpretados y los usuarios actúan como si “grado de confianza” fuera sinónimo de uniformidad dentro del intervalo.
Valores p iguales no proporcionan igual evedencia acerca de la hipótesis, Harrel Jr., F. E. (2000)

Distribución Exponencial: Familia Conjugada

La distribución exponencial tiene función de densidad de probabilidad dada por

$$f(x) = \theta e^{-\theta x} \qquad x \in (0, \infty)$$

(Ojo: hay que tener cuidado con la forma de parametrizar!)

- Suponga que $X_1, \cdots, X_n$ es una muestra de una distribución exponencial con parámetro desconocido θ.
- Supongamos que la distribución apriori de θ es una gamma con parámetros $\alpha(>0)$ y $\beta(>0)$.
- La distribución posterior de θ cuando $X_i = x_i$, para $i = 1, \cdots, n$ es una gamma con parámetros $\alpha+n$ y $\beta+\sum_{i=1}^{n} x_i$.

Distribución Exponencial: Intervalo de credibilidad

El intervalo de credilidad de probabilidad $(1-\gamma)100\,\%$ es

$$(LI; LS)$$

donde

$$\frac{\gamma}{2} = \int_0^{LI} Gamma\left(\alpha + n; \beta + \sum_{i=1}^{n} x_i\right) d\theta$$

y

$$1 - \frac{\gamma}{2} = \int_0^{LS} Gamma\left(\alpha + n; \beta + \sum_{i=1}^{n} x_i\right) d\theta$$

Distribución Exponencial: Intervalo de credibilidad cuando la apriori es Laplace

El intervalo de credilidad de probabilidad $(1-\gamma)100\,\%$ es

$$(LI; LS)$$

donde

$$\frac{\gamma}{2} = \int_0^{LI} Gamma\left(n; \sum_{i=1}^{n} x_i\right) d\theta$$

y

$$1 - \frac{\gamma}{2} = \int_0^{LS} Gamma\left(n; \sum_{i=1}^{n} x_i\right) d\theta$$

Página www

Página de Abertura

Contenido

Página 12 de 56

Regresar

Full Screen

Cerrar

Abandonar

```
> tiempo.hasta.primer.gol.2001<-scan()
1: 57 26 50 67 55 27 8 61 26 35 8 54 35 14 72 12 31 27 7 4 63 24 4 8
27: 31 27 7 4 63 24 4 63 24 4 82 8 7 31 15 2 49 85 38 12 24 41 64 47
54:
Read 53 items
> y<-sum(tiempo.hasta.primer.gol.2001)
> n<-length(tiempo.hasta.primer.gol.2001)
> n
[1] 53
> y
[1] 1727
> qgamma(0.025,n,rate=y)
[1] 0.02298821
> qgamma(0.975,n,rate=y)
[1] 0.03948528
> 1/qgamma(0.975,n,rate=y)
[1] 25.32589
> 1/qgamma(0.025,n,rate=y)
[1] 43.50056
> m<-mean(tiempo.hasta.primer.gol.2001)
> s<-sd(tiempo.hasta.primer.gol.2001)
> m-1.96*s/sqrt(n)
[1] 25.74058
> m+1.96*s/sqrt(n)
[1] 39.42923
```

Ejemplo: Distribución Uniforme

Rossman et al. (1998) presentan la contrucción de la región de mayor probabilidad para el "parámetro" de la distribución uniforme $U(0, \theta)$. La estadística clásica nos presenta, asumiendo que $X_1, \cdots, X_n$ sea una muestra aleatoria,

Estimador de Máxima Verosimilitud	$\max\{X_i\}$
Estimador de Mínima Varianza Insesgado	$\frac{n+1}{n}\max\{X_i\}$

Si escogemos una distribución apriori impropia o aplanada de la forma $\xi(\theta) = 1$ para $\theta > 0$, la distribución posterior es proporcional a la función de verosimilitud,

$$\xi(\theta|X) \propto \frac{1}{\theta^n} \text{ para } \theta \geq \max\{X_i\}$$

La constante de proporcionalidad, que vuelve la distribución posterior propia es $(n-1)\left(\max\{X_i\}\right)^{n-1}$. Bajo la función de pérdida cuadrática el estimador bayesiano es igual a la media aposteriori

$$E[\theta|X] = \int_{-\infty}^{\infty} \theta \cdot \xi\left(\theta|X\right)\, d\theta = \frac{n-1}{n-2}\max\{X_i\}$$

Un intervalo de probabilidad del 95 % se halla resolviendo

$$\int_{LI}^{LS} \frac{(n-1)\left(\max\{X_i\}\right)^{n-1}}{\theta^n}\, d\theta$$

Región de la Densidad Posterior Más Alta (RDPMA)

Si $p(\theta|Y)$ denota la densidad posterior entonces podemos definir un intervalo de credibilidad utilizando la RDPMA.

Definición

(Box y Tiao, 1973) Una región R en un espacio parametral Θ es llamada la región de la densidad posterior más alta (RDPMA) de contenido α si

1. $P(\theta \in R|Y) = \alpha$
2. Para $\theta_1 \in R$ y $\theta_2 \notin R$, se cumple $P(\theta_1 \in R|Y) \geq P(\theta_2 \in R|Y)$.

Para un contenido de probabilidad α, la RDPMA tiene el volumen más pequeño en el espacio parametral.

Ejemplo

Vamos a construir un intervalo del 95 % de probabilidad de la mayor densidad para el parámetro de la Poisson vía simulación asumiendo que la aposteriori es $Gamma(17, 5)$

en R: librería `hdrcde`

La función `hdr()` permite calcular un intervalo de probabilidad de alta densidad.

```
library(hdrcde)
x<-rgamma(10000,17,rate=5)
hdr(x, prob = c(50, 95, 99))

$hdr
        [,1]     [,2]
99% 1.493444 5.751071
95% 1.864606 5.063168
50% 2.710668 3.776178

$mode
[1] 3.211995

$falpha
        1%         5%        50%
0.02187424 0.08048806 0.38669267
```

Mientras que el intervalo bayesiano hallado con ambas colas iguales a $\alpha/2$ es

```
> qgamma(0.025,17,rate=5)
[1] 1.980625

> qgamma(0.975,17,rate=5)
[1] 5.1966
```

En el caso de los goles del local

```
library(hdrcde)
> x<-rgamma(10000,390,rate=251)
> hdr(x, prob = c(50, 95, 99))
$hdr
        [,1]     [,2]
99% 1.359219 1.759533
95% 1.402750 1.710359
50% 1.498031 1.604939

$mode
[1] 1.54933

$falpha
       1%        5%       50%
0.2376944 0.8444804 3.8998675

> qgamma(0.025,390,rate=251)
[1] 1.403384
> qgamma(0.975,390,rate=251)
[1] 1.711731
```

Pruebas de Hipótesis

Ejemplo: Poderes Sobrenaturales

Bayarri y Berger en la reunión anual que se lleva a cabo en Valencia (España) presentaron el siguiente caso de sicokinesis: Tres investigadores (Schmidt, Jahn y Radin) en 1987 utilizaron un generador cuántico que recibe una fila de partículas y él desvía cada partícula, independientemente de las otras, hacia una luz roja o una luz verde con igual probabilidad. Se le pidió a un sujeto quien alegaba tener poderes sicokinéticos que tratara de influenciar el generador de tal forma que las partículas se fueran para la luz roja. Se generaron 104.490.000 partículas y se contaron 52.263.470 partículas que se fueron hacia la luz roja. Habrá suficiente evidencia que permita decir que el sujeto tiene poderes sicokinéticos?

Podemos pensar en este exprimento así: Cada partícula corresponde a un ensayo $Bernoulli(\pi)$, y un éxito será si la partícula se va para la luz roja. Si X denota el número de éxitos, $X \sim Binomial\,(n, \pi)$. Tenemos $x = 52{,}263{,}470$ como la observación real. Se necesita probar

$$
\begin{aligned}
H_0 : \pi &= \frac{1}{2} \quad \text{(El sujeto no tiene poderes)}\\
H_1 : \pi &\neq \frac{1}{2} \quad \text{(El sujeto tiene poderes)}
\end{aligned}
$$

El $valor - p = P_{H_0}\left(\left|X - \frac{n}{2}\right| \geq \left|x - \frac{n}{2}\right|\right) \approx 0{,}0003$ nos lleva a concluir que hay una fuerte evidencia contra H_0.

Si pensamos bayesianamente necesitamos una distribución apriori, pero ahora definida sobre las hipótesis en juego:

$$\xi\left(H_i\right) = \text{probabilidad apriori de que } H_i \text{ sea cierta, } i = 0, 1.$$

Bajo $H_1 : \pi \neq 1/2$, sea $\xi(\pi)$ la densidad apriori sobre π. El Bayes objetivo selecciona

$$Pr\left(H_0\right) = Pr\left(H_1\right) = \frac{1}{2}$$

con $\xi(\pi) = 1 \;\; (0 < \pi < 1)$

La probabilidad posterior de la hipótesis

$$\begin{aligned} Pr\left(H_0|x\right) &= \text{probabilidad de que} H_0 \text{ sea cierta dados los datos } x \\ &= \frac{f\left(x|\pi=1/2\right)Pr\left(H_0\right)}{Pr\left(H_0\right)f\left(x|\pi=1/2\right)+Pr\left(H_1\right)\int f\left(x|\pi\right)\xi(\pi)\,d\pi} \end{aligned}$$

Para la apriori objetiva

$$Pr\left(H_0|x = 52{,}263{,}470\right) \approx 0{,}92$$

La densidad posterior en $H_1 : \pi \neq 1/2$ es

$$\xi\left(\pi|x, H_1\right) \propto \xi(\pi)f(x|\pi) \propto 1 \times \pi^x(1-\pi)^{n-x},$$

que es una $Beta\left(52{,}263{,}470,\ \ 52{,}226{,}530\right)$

Lanzamiento de un par de dados

En un juego de parqués se registraron los resultados del lanzamiento de un par de dados 130 veces. A partir de estos resultados quiere uno ver si los dados son conjuntamente buenos.

Resultado	2	3	4	5	6	7	8	9	10	11	12
Frecuencia	4	8	10	11	22	14	22	18	10	5	6

Nos podemos preguntar si con los datos anteriores podríamos jugar tranquilamente este juego de parqués, o sea si los dados son buenos o están cargados.
Si el par de dados fueran perfectos, entonces el modelo teórico sería el que aparece en la siguiente tabla:

Resultado	2	3	4	5	6	7	8	9	10	11	12
Probabilidad	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

En los 130 lanzamientos de los dados esperaríamos hallar

Resultado	2	3	4	5	6	7
Esperada	3.61	7.22	10.83	14.44	18.06	21.67
Resultado	8	9	10	11	12	
Esperada	18.06	14.44	10.83	7.22	3.61	

Suponga las dos hipótesis

$$\begin{aligned} H_0 &: \text{ Los dados están buenos} \\ H_1 &: \text{ Los dados están sesgados} \end{aligned}$$

Suponga que apriori no tenemos información que nos haga dudar sobre la calidad de los dados y escogemos

$$\xi(H_0) = 0{,}9$$
$$\xi(H_1) = 0{,}1$$