

## **CAPÍTULO 3. MODELOS GAUSSIANOS**

### **Para leer**

Gelman et al (1995) Capítulo 3, Secciones 3.1 – 3.4, 3.6 y 3.8.

Box y Tiao (1973), Capítulo 2, Secciones 2.1 – 2.5.

Lee (1997), Capítulo 2, Capítulo 5, Secciones 5.1–5.4.

En este capítulo, se estudian problemas de una y dos muestras con datos normales.

## Problemas de una muestra

Sea  $X|\mu, \sigma^2 \sim \mathcal{N}(\mu, \sigma^2)$ . Dada una muestra de datos,  $\mathbf{x}$ , se quiere hacer inferencia sobre  $\mu$  y/o  $\sigma^2$ .

Recordamos la verosimilitud de  $\mu$  y  $\sigma^2$  del Ejemplo 21:

$$l(\mu, \sigma^2|\mathbf{x}) \propto \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2} [(n-1)s^2 + n(\mu - \bar{x})^2]\right).$$

Se quiere definir una distribución a priori conjugada para  $\mu, \sigma^2$ .

**Observación 26** *A menudo en la inferencia bayesiana, se transforma  $\phi = 1/\sigma^2$ . Por tanto  $\phi$  es la precisión. Se tiene*

$$l(\mu, \phi|\mathbf{x}) \propto \phi^{n/2} \exp\left(-\frac{\phi}{2} [(n-1)s^2 + n(\mu - \bar{x})^2]\right)$$

## Una distribución a priori conjugada

Se ha visto anteriormente que dado  $\phi$ , la distribución a priori conjugada para  $\mu$  es normal. Considerando la verosimilitud en términos de  $\phi$ , parece a una densidad gamma.

Se descompone la distribución a priori como

$$f(\mu, \phi) = f(\mu|\phi)f(\phi)$$

y se define

$$\begin{aligned}\mu|\phi &\sim \mathcal{N}\left(m, \frac{1}{\alpha\phi}\right) \\ \phi &\sim \mathcal{G}\left(\frac{a}{2}, \frac{b}{2}\right)\end{aligned}$$

Entonces, la distribución conjunta es

$$\begin{aligned}f(\mu, \phi) &\propto \phi^{\frac{1}{2}} \exp\left(-\frac{\alpha\phi}{2}(\mu - m)^2\right) \phi^{\frac{a}{2}-1} \exp\left(-\frac{b\phi}{2}\right) \\ &\propto \phi^{\frac{a+1}{2}-1} \exp\left(-\frac{\phi}{2}\left[b + \alpha(\mu - m)^2\right]\right)\end{aligned}$$

**Observación 27** La distribución conjunta de  $\mu$  y  $\phi$  se llama la **distribución normal gamma** con parámetros  $m$ ,  $\alpha$ ,  $a/2$  y  $b/2$ .

$$\mu, \phi \sim \mathcal{NG}(m, \alpha, a/2, b/2)$$

Es interesante explorar las propiedades de esta distribución. En primer lugar, ¿cuál es la distribución de  $\sigma^2$ ?

**Teorema 4 1.** La distribución inducida de  $\tau = \sigma^2$  es una distribución **gamma invertida**

$$f(\tau) = \frac{\left(\frac{b}{2}\right)^{\frac{a}{2}}}{\Gamma\left(\frac{a}{2}\right)} \tau^{-(\frac{a}{2}+1)} e^{-\frac{b}{2\tau}}$$

para  $\tau > 0$ .

2. Un intervalo de credibilidad de 95 % para  $\sigma^2$  es

$$\left( \frac{b}{\chi_a^2(0,975)}, \frac{b}{\chi_a^2(0,025)} \right).$$

## Demostración

1. Siendo  $\tau = \frac{1}{\phi}$ , luego  $\phi = \frac{1}{\tau}$  y  $\frac{d\phi}{d\tau} = -\frac{1}{\tau^2}$  y utilizando la fórmula para transformación de variables,

$$f(\tau) = f_{\phi} \left( \frac{1}{\tau} \right) \left| \frac{d\phi}{d\tau} \right|$$

se tiene el resultado.

2. Sea  $c$  el punto  $0,975 \times 100\%$  de la distribución de  $\sigma^2$ . Luego:

$$0,975 = P(\sigma^2 < c) = P\left(\phi > \frac{1}{c}\right) = P\left(b\phi > \frac{b}{c}\right)$$

$$\text{y luego } P\left(b\phi < \frac{b}{c}\right) = 0,025.$$

Sabiendo que  $b\phi \sim \chi_a^2$ , entonces se tiene  $\frac{b}{c} = \chi_a^2(0,025)$  y  $c = \frac{b}{\chi_a^2(0,025)}$ .

◇

## La distribución marginal de $\mu$ .

**Teorema 5**  $\mu \sim \mathcal{T}(a, m, b/(\alpha a))$ , una distribución *t* de Student (escalada y no centrada). Un intervalo de credibilidad de 95 % para  $\mu$  es

$$m \pm \sqrt{b/(a\alpha)} t_{a, .025}$$

### Demostración

$$\begin{aligned} f(\mu) &= \int_0^\infty f(\mu, \phi) d\phi \\ &\propto \int_0^\infty \phi^{\frac{a+1}{2}-1} \exp\left(-\frac{\phi}{2} (b + \alpha(\mu - m)^2)\right) d\phi \\ &\propto (b + \alpha(\mu - m)^2)^{-\left(\frac{a+1}{2}\right)} \\ &\propto \left(1 + \frac{1}{a} \left(\frac{\mu - m}{\sqrt{b/a\alpha}}\right)^2\right)^{-\left(\frac{a+1}{2}\right)} \end{aligned}$$

el núcleo de una distribución *t* escalada y no centrada con  $a$  grados de libertad, y media  $m$ .

◇

**Observación 28** Transformando  $\theta = \frac{\mu - m}{\sqrt{b/a\alpha}}$ , se tiene  $\theta \sim t_a$ , una distribución  $t$  estandar.

## La distribución predictiva de $X$

Supongamos que se quiere calcular la distribución predictiva de  $X$ .

**Teorema 6** Sea  $X|\mu, \phi \sim \mathcal{N}(\mu, 1/\phi)$  y  $\mu, \phi \sim \mathcal{NG}(m, \alpha, a/2, b/2)$ . La distribución marginal de  $X$  es

$$X \sim \mathcal{T}(a, m, b(\alpha + 1)/(\alpha a)).$$

Un intervalo predictivo de 95 % para  $X$  es

$$m \pm \sqrt{\frac{b(\alpha + 1)}{\alpha a}} t_a(,025).$$

## Demostración

Sea  $X = \mu + \epsilon$  donde  $\mu|\phi \sim \mathcal{N}(m, 1/(\alpha\phi))$  y  $\epsilon|\phi \sim \mathcal{N}(0, 1/\phi)$ . Obviamente, se tiene  $X|\mu, \phi \sim \mathcal{N}(\mu, 1/\phi)$  y además, mediante las propiedades de la distribución normal,

$$X|\phi \sim \mathcal{N}\left(m, \frac{1}{\phi} \left(1 + \frac{1}{\alpha}\right)\right)$$

Defina  $\alpha' = \left(1 + \frac{1}{\alpha}\right)^{-1} = \frac{\alpha}{\alpha+1}$  y entonces se tiene  $X, \phi \sim \mathcal{NG}(m, \alpha', a/2, b/2)$ .

Mediante el teorema 5, se tiene

$$X \sim \mathcal{T}(a, m, b/(\alpha'a))$$

y sustituyendo para  $\alpha'$  se demuestra el resultado.

◇

## La distribución a posteriori conjunta

Multiplicando por la verosimilitud (ver Observación 26) se demuestra el siguiente teorema:

**Teorema 7** *La distribución a posteriori es*

$$\mu, \phi | \mathbf{x} \sim \mathcal{NG}(m^*, \alpha^*, a^*/2, b^*/2)$$

donde

$$\begin{aligned} a^* &= a + n \\ b^* &= b + (n - 1)s^2 + \frac{\alpha n}{\alpha + n}(m - \bar{x})^2 \\ \alpha^* &= \alpha + n \\ m^* &= \frac{\alpha m + n\bar{x}}{\alpha + n} \end{aligned}$$

Luego

$$\begin{aligned} \mu | \phi, \mathbf{x} &\sim \mathcal{N}\left(m^*, \frac{1}{\alpha^* \phi}\right), \\ \phi | \mathbf{x} &\sim \mathcal{G}\left(\frac{a^*}{2}, \frac{b^*}{2}\right), \\ \mu | \mathbf{x} &\sim \mathcal{T}(a^*, m^*, b^*/(\alpha^* a^*)), \\ X | \mathbf{x} &\sim \mathcal{T}(a^*, m^*, b^*(\alpha^* + 1)/(\alpha^* a^*)). \end{aligned}$$

## Demostración

$$\begin{aligned}
 f(\mu, \phi | \mathbf{x}) &\propto l(\mu, \phi | \mathbf{x}) f(\mu, \phi) \\
 &\propto \phi^{\frac{n}{2}} \exp\left(-\frac{\phi}{2} [(n-1)s^2 + n(\mu - \bar{x})^2]\right) \times \\
 &\quad \phi^{\frac{a+1}{2}-1} \exp\left(-\frac{\phi}{2} [b + \alpha(\mu - m)^2]\right) \\
 &\propto \phi^{\frac{a+n+1}{2}-1} \exp\left(-\frac{\phi}{2} [b + \alpha(\mu - m)^2 + (n-1)s^2 + \right. \\
 &\quad \left. n(\mu - \bar{x})^2]\right) \\
 &\propto \phi^{\frac{a^*+1}{2}-1} \exp\left(-\frac{\phi}{2} [b + (n-1)s^2 + (\alpha + n)\mu^2 - \right. \\
 &\quad \left. 2\mu(\alpha m + n\bar{x}) + \alpha m^2 + n\bar{x}^2]\right) \\
 &\propto \phi^{\frac{a^*+1}{2}-1} \exp\left(-\frac{\phi}{2} [b + (n-1)s^2 + \alpha^*(\mu - m^*)^2 + \right. \\
 &\quad \left. \alpha m^2 + n\bar{x}^2 - \alpha^* m^{*2}]\right)
 \end{aligned}$$

$$\begin{aligned}
&\propto \phi^{\frac{a^*+1}{2}-1} \exp \left( -\frac{\phi}{2} \left[ b + (n-1)s^2 + \alpha^*(\mu - m^*)^2 + \right. \right. \\
&\quad \left. \left. \alpha m^2 + n\bar{x}^2 - \frac{(\alpha m + n\bar{x})^2}{\alpha + n} \right] \right) \\
&\propto \phi^{\frac{a^*+1}{2}-1} \exp \left( -\frac{\phi}{2} \left[ b + (n-1)s^2 + \alpha^*(\mu - m^*)^2 + \right. \right. \\
&\quad \left. \left. \frac{\alpha n}{\alpha + n} (m - \bar{x})^2 \right] \right) \\
&\propto \phi^{\frac{a^*+1}{2}-1} \exp \left( -\frac{\phi}{2} [b^* + \alpha^*(\mu - m^*)^2] \right) \\
\mu, \phi | \mathbf{x} &\sim \mathcal{NG}(m^*, \alpha^*, a^*/2, b^*/2)
\end{aligned}$$

*Los resultados para la distribución marginal de  $\mu$  y la distribución predictiva de una nueva observación  $X$  siguen de los teoremas 5 y 6.*

◇

**Ejemplo 26** Suponer que  $X|\mu, \phi \sim \mathcal{N}(\mu, 1/\phi)$  con las distribuciones a priori  $\mu|\phi \sim \mathcal{N}(85, 1/\phi)$  y  $\phi \sim \mathcal{G}\left(\frac{4}{2}, \frac{350}{2}\right)$ .

Dada una muestra con  $n = 100$ ,  $\bar{x} = 89$  y  $s^2 = 30$ , calcular la distribución a posteriori de  $\mu$  y  $\phi$ .

Tenemos

$$a^* = 4 + 100 = 104$$

$$\begin{aligned} b^* &= 350 + 99 \times 30 + \frac{1 \times 100}{1 + 100} (85 - 89)^2 \\ &= 3335,84 \end{aligned}$$

$$\alpha^* = 1 + 100 = 101$$

$$\begin{aligned} m^* &= \frac{1 \times 85 + 100 \times 89}{1 + 100} \\ &= 88,96 \end{aligned}$$

*Entonces*

$$\mu, \phi | \mathbf{x} \sim \mathcal{NG} \left( 88,96, 101, \frac{104}{2}, \frac{3335,84}{2} \right)$$

$$\mu | \phi, \mathbf{x} \sim \mathcal{N} \left( 88,96, \frac{1}{101\phi} \right) \quad y$$

$$\phi | \mathbf{x} \sim \mathcal{G} \left( \frac{104}{2}, \frac{3335,84}{2} \right).$$

*También calculamos intervalos de credibilidad para  $\mu$  antes y después de ver los datos.*

*Antes de ver los datos tenemos el intervalo*

$$85 \pm \sqrt{\frac{335}{4 \times 1}} t_4(,025) \approx (59,6, 110,4)$$

*Después de ver los datos tenemos el intervalo*

$$88,96 \pm \sqrt{\frac{3335,84}{104 \times 101}} t_{104}(,025) \approx (87,84, 90,08)$$

*Finalmente sacamos intervalos predictivos para una observación  $X$  antes y después de ver los datos.*

*Antes de ver los datos tenemos el intervalo*

$$85 \pm \sqrt{\frac{335 \times (4 + 1)}{4 \times 1}} t_4(,025) \approx (28,19, 141,81)$$

*Después de ver los datos tenemos*

$$88,96 \pm \sqrt{\frac{3335,84 \times (104 + 1)}{104 \times 101}} t_{104}(,025) \\ \approx (77,53, 100,39)$$

*Los intervalos predictivos son mucho más anchos que los intervalos de credibilidad para la media.*

## Resultados equivalentes a resultados clásicos

Dada la distribución a priori *impropia*

$$f(\mu, \phi) \propto \frac{1}{\phi},$$

la distribución a posteriori es

$$\begin{aligned} f(\mu, \phi | \mathbf{x}) &\propto \phi^{\frac{n}{2}-1} \exp\left(-\frac{\phi}{2} [(n-1)s^2 + n(\mu - \bar{x})^2]\right) \\ &\propto \phi^{\frac{(n-1)+1}{2}-1} \exp\left(-\frac{\phi}{2} [(n-1)s^2 + \right. \\ &\quad \left. + n(\mu - \bar{x})^2]\right) \end{aligned}$$

que es el núcleo de una distribución normal gamma.

$$\mu, \phi | \mathbf{x} \sim \mathcal{NG}\left(\bar{x}, n, \frac{n-1}{2}, \frac{(n-1)s^2}{2}\right).$$

Luego  $\mu | \phi, \mathbf{x} \sim \mathcal{N}\left(\bar{x}, \frac{1}{n\phi}\right)$  y  $\phi | \mathbf{x} \sim \mathcal{G}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$  y por tanto  $\mu | \mathbf{x} \sim \mathcal{T}(n-1, \bar{x}, s^2/n)$ .

Un intervalo de credibilidad bayesiano de 95 % para  $\mu$  es

$$\bar{x} \pm t_{n-1}(,025) \frac{s}{\sqrt{n}}$$

que coincide con el intervalo clásico de confianza.

La distribución predictiva para una nueva observación  $X$  es  $X \sim \mathcal{T}(n-1, \bar{x}, (1+1/n)s^2)$ . Un intervalo predictivo de 95 % para  $X$  es

$$\bar{x} \pm t_{n-1}(,025) s \sqrt{1 + \frac{1}{n}}.$$

Finalmente, usando Teorema 4, un intervalo de credibilidad de 95 % para  $\sigma^2$  es

$$\left( \frac{s^2}{\chi_{n-1}^2(0,975)}, \frac{s^2}{\chi_{n-1}^2(0,025)} \right)$$

igual al intervalo clásico de confianza.

## Problemas de dos muestras

Se consideran varios problemas tanto en la literatura clásica como bayesiana:

1. Muestras dependientes: datos apareados.
2. Muestras independientes:
  - a) con varianzas conocidas,
  - b) con varianzas desconocidas pero iguales,
  - c) con varianzas desconocidas no iguales.
3. La razón de dos varianzas.

Sólo se considera en detalle los problemas con datos no apareados y varianzas desconocidas. Se resume los otros problemas.

## **Datos apareados**

Igual a la inferencia clásica, tomando las diferencias de cada pareja de de datos, el problema se reduce a un problema de una muestra.

## **Datos independientes con varianzas conocidas**

Dadas distribuciones a priori normales para las medias desconocidas, y usando el Ejemplo 21, las distribuciones a posteriori son también normales y luego la diferencia entre las medias es normal. Distribuciones a priori uniformes para las dos medias implica que la inferencia bayesian coincide numéricamente con la inferencia clásica.

## Datos independientes, con varianzas iguales pero desconocidas

Sea  $X_i|\mu_1, \phi \sim \mathcal{N}(\mu_i, 1/\phi)$  con muestras de tamaño  $n_i$ , para  $i = 1, 2$ . Dada la distribución a priori no informativa

$$f(\mu_1, \mu_2, \phi) \propto \frac{1}{\phi},$$

se tiene

$$f(\mu_1, \mu_2, \phi | \mathbf{x}_1, \mathbf{x}_2) \propto \phi^{\frac{n_1+n_2-1}{2}} \exp\left(-\frac{\phi}{2} \left[ \sum_{i=1}^2 (n_i - 1) s_i^2 + n_i (\mu_i - \bar{x}_i)^2 \right]\right)$$
$$\mu_i | \phi, \mathbf{x}_i \sim \mathcal{N}\left(\bar{x}_i, \frac{1}{n_i \phi}\right) \quad \text{para } i = 1, 2.$$

Integrando por  $\mu_1$  y  $\mu_2$  sucesivamente, se ve que la distribución marginal de  $\phi$  es

$$\phi | \mathbf{x} \sim \mathcal{G}\left(\frac{n_1 + n_2 - 2}{2}, \sum_{i=1}^2 \frac{(n_i - 1) s_i^2}{2}\right).$$

Además, definiendo  $\delta = \mu_1 - \mu_2$ , se tiene

$$\delta | \phi, \mathbf{x}_1, \mathbf{x}_2 \sim \mathcal{N} \left( \bar{x}_1 - \bar{x}_2, \frac{1}{\phi} \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \right)$$

y luego,  $\delta, \phi$  tiene una distribución normal gamma

$$\delta, \phi \sim \mathcal{NG} \left( \bar{x}_1 - \bar{x}_2, \left( \frac{1}{n_1} + \frac{1}{n_2} \right)^{-1}, \frac{n_1 + n_2 - 2}{2}, \sum_{i=1}^2 \frac{(n_i - 1)s_i^2}{2} \right)$$

Un intervalo de credibilidad de 95 % para  $\delta$  es

$$\bar{x}_1 - \bar{x}_2 \pm s_{comb} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}$$

donde  $s_{comb}^2 = \frac{\sum_{i=1}^2 (n_i - 1)s_i^2}{n_1 + n_2 - 2}$ .

El intervalo es igual al intervalo frecuentista.

## Varianzas desconocidas y no iguales: el problema de Behrens y Fisher

Se tienen dos poblaciones  $X_1 \sim \mathcal{N}(\mu_1, 1/\phi_1)$  y  $X_2 \sim \mathcal{N}(\mu_2, 1/\phi_2)$  y se quiere hacer inferencia sobre  $\delta = \mu_1 - \mu_2$ .

Se observan dos muestras  $(n_1, \bar{x}_1, s_1^2)$  y  $(n_2, \bar{x}_2, s_2^2)$ .

Pongamos las distribuciones a priori no informativas

$$f(\mu_i \phi_i) \propto \frac{1}{\phi_i} \quad i = 1, 2.$$

Sabemos que las distribuciones a posteriori de  $\mu_1$  y  $\mu_2$  son

$$\begin{aligned} \mu_1 | \mathbf{x}_1 &\sim \mathcal{T}(n_1 - 1, \bar{x}_1, s_1^2/n_1) \\ \mu_2 | \mathbf{x}_2 &\sim \mathcal{T}(n_2 - 1, \bar{x}_2, s_2^2/n_2) \end{aligned}$$

La distribución a posteriori de  $\delta$  es la distribución de la diferencia entre dos variables  $t$  de Student.

Se puede intentar hallar la distribución de  $\delta$  a través de la fórmula para la diferencia de dos variables:

$$f(\delta|\mathbf{x}_1, \mathbf{x}_2) = \int f_{\mu_1}(\delta + \mu_2|\mathbf{x}_1) f(\mu_2|\mathbf{x}_2) d\mu_2$$

pero es mas fácil considerar la función *normalizada*

$$\delta' = \frac{\delta - (\bar{x}_1 - \bar{x}_2)}{(s_1^2/n_1 + s_2^2/n_2)^{1/2}}.$$

**Teorema 8** Siendo  $\tan \omega = \frac{s_1/\sqrt{n_1}}{s_2/\sqrt{n_2}}$ , entonces

$$\delta' = T_1 \sin \omega - T_2 \cos \omega$$

donde  $T_1 = \frac{\mu_1 - \bar{x}_1}{s_1/\sqrt{n_1}}$  and  $T_2 = \frac{\mu_2 - \bar{x}_2}{s_2/\sqrt{n_2}}$  tienen distribuciones de  $t$  estandares con  $n_1 - 1$  y  $n_2 - 1$  grados de libertad respetivamente.

## Demostración

$$\begin{aligned} \delta' &= \frac{(\mu_1 - \mu_2) - (\bar{x}_1 - \bar{x}_2)}{(s_1^2/n_1 + s_2^2/n_2)^{1/2}} \\ &= \frac{(\mu_1 - \bar{x}_1) - (\mu_2 - \bar{x}_2)}{(s_1^2/n_1 + s_2^2/n_2)^{1/2}} \\ &= \frac{s_1/\sqrt{n_1} \frac{\mu_1 - \bar{x}_1}{s_1/\sqrt{n_1}} - s_2/\sqrt{n_2} \frac{\mu_2 - \bar{x}_2}{s_2/\sqrt{n_2}}}{(s_1^2/n_1 + s_2^2/n_2)^{1/2}} \\ &= \frac{s_1/\sqrt{n_1}}{(s_1^2/n_1 + s_2^2/n_2)^{1/2}} T_1 - \frac{s_2/\sqrt{n_2}}{(s_1^2/n_1 + s_2^2/n_2)^{1/2}} T_2 \end{aligned}$$

Cuadrando y sumando los dos pesos, el resultado es 1 y entonces

$$\delta' = T_1 \sin \omega - T_2 \cos \omega.$$

◇

**Definición 5** *La distribución de*

$$T_{\nu_1} \sin \omega - T_{\nu_2} \cos \omega$$

*donde  $T_\nu$  es una variable  $t$  de Student con  $\nu$  grados de libertad, se llama **la distribución de Behrens y Fisher** (Behrens 1929) con parámetros  $\nu_1, \nu_2$  y  $\omega$ :  $\mathcal{BF}(\nu_1, \nu_2, \omega)$ .*

Entonces se concluye que

$$\delta' \sim \mathcal{BF} \left( n_1 - 1, n_2 - 1, \tan^{-1} \frac{s_1 / \sqrt{n_1}}{s_2 / \sqrt{n_2}} \right).$$

**Observación 29** *Esta solución bayesiana del problema equivale numéricamente a la solución fiducial propuesta por Fisher (1939).*

## Cálculo de intervalos de credibilidad

Se puede calcular intervalos aproximados a través de la siguiente aproximación debida a Patil (1965).

Si  $X \sim \mathcal{BF}(\nu_1, \nu_2, \omega)$  entonces  $X/a \approx t_b$  donde

$$\begin{aligned}f_1 &= \left( \frac{n_1 - 1}{n_1 - 3} \right) \sin^2 \omega + \left( \frac{n_2 - 1}{n_2 - 3} \right) \cos^2 \omega \\f_2 &= \frac{(n_1 - 1)^2}{(n_1 - 3)^2(n_1 - 5)} \sin^4 \omega + \\&\quad \frac{(n_2 - 1)^2}{(n_2 - 3)^2(n_2 - 5)} \cos^4 \omega \\b &= 4 + (f_1^2 / f_2) \\a &= \sqrt{f_1(b - 2) / b}\end{aligned}$$

**Observación 30** *Otro método sencillo es muestrear las distribuciones a posteriori de  $\mu_1$  y  $\mu_2$  para simular una muestra de la distribución de  $\delta$ .*

## La solución (?) clásica

No existen resultados parecidos usando inferencia clásica porque, desde el punto de vista clásico, la distribución de muestreo de  $\delta'$  depende de la razón  $\phi_1/\phi_2$  de las dos precisiones desconocidas.

La solución típica es aproximar la distribución de muestreo con una distribución t:

$$\delta' \approx t_r \quad \text{donde} \quad r = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\left(\frac{s_1^4/n_1^2}{n_1-1} + \frac{s_2^4/n_2^2}{n_2-1}\right)}.$$

El problema es que no se puede evaluar la calidad de la aproximación para una muestra dada.

**Ejemplo 27** *Se toman dos muestras de tamaños 16 y 9 con  $(\bar{x}, s_1^2) = (1, 4)$  y  $(\bar{y}, s_2^2) = (2, 9)$ .*

*Suponiendo distribuciones a priori no informativas,  $\delta' \sim \mathcal{BF}(15, 8, \omega)$  donde  $\omega = \tan^{-1}(0,5)$  y un intervalo de credibilidad de 95% (utilizando la aproximación de Patil que implica  $a = 1,0164, b = 9,819$ ) es  $(-3,5384, 1,5384)$*

*El intervalo de confianza clásico (calculado usando  $r = 12,0968$ ) es  $(-3,4338, 1,4338)$  un poco distinto.*

## Inferencia para la razón de dos varianzas

Supongamos que se quiere hacer inferencia sobre la razón de las dos varianzas

$$\zeta = \sigma_1^2 / \sigma_2^2 = \phi_2 / \phi_1.$$

Dadas dos muestras  $(n_i, \bar{x}_i, s_i^2)$ , para  $i = 1, 2$ , como anteriormente y distribuciones a priori no informativas  $f(\mu_i, \phi_i) \propto 1/\phi_i$ , las distribuciones a posteriori de las precisiones son gamma

$$\phi_i | \text{data} \sim \mathcal{G} \left( \frac{n_i - 1}{2}, \frac{(n_i - 1)s_i^2}{2} \right)$$

y luego  $(n_i - 1)s_i^2 \phi_i \sim \chi_{n_i - 1}^2$ . Entonces recordando que una variable  $F$  es el resultado de dividir una variable ji-cuadrado partidas por sus precisiones, se tiene

$$\frac{s_2^2}{s_1^2} \zeta | \text{data} \sim \mathcal{F}(n_2 - 1, n_1 - 1).$$

Un intervalo de credibilidad de 95 % para  $\zeta$  es

$$\left( \frac{s_1^2}{s_2^2} F_{n_1-1}^{n_2-1}(0,025), \frac{s_1^2}{s_2^2} F_{n_1-1}^{n_2-1}(0,975) \right)$$

que coincide con el intervalo clásico.

**Observación 31** *Es fácil extender el análisis de este problema y al problema de Behrens y Fisher al caso en que se utiliza una distribución a priori normal gamma.*

## Un problema y una solución

Volvemos al problema de la inferencia para la media  $\mu$  de la población donde la precisión  $\phi$  es desconocida.

La elección de distribuciones a priori conjugadas es restrictiva y no muy natural porque se supone que la distribución de la media depende del valor de la varianza poblacional.

En muchos problemas es más razonable suponer distribuciones a priori independientes para la media y la varianza. Pero entonces caso la inferencia no es conjugada.

**Ejemplo 28** Sea  $X|\mu, \phi \sim \mathcal{N}(\mu, 1/\phi)$  con distribuciones a priori  $\mu \sim \mathcal{N}(m, 1/\delta)$  y  $\phi \sim \mathcal{G}(a/2, b/2)$ .

*Dados los datos  $\mathbf{x}$ , no se puede calcular la distribución a posteriori de  $\mu$  en una forma cerrada. Pero es posible calcular las distribuciones acondicionadas.*

$$\mu|\phi, \mathbf{x} \sim \mathcal{N}\left(\frac{\delta m + n\phi\bar{x}}{\delta + n\phi}, \frac{1}{\delta + n\phi}\right)$$

$$\phi|\mu, \mathbf{x} \sim \mathcal{G}\left(\frac{a + n}{2}, \frac{b + (n - 1)s^2 + n(\mu - \bar{x})^2}{2}\right)$$

*Ahora, se puede emplear el muestreo Gibbs (Gelfand y Smith 1990) para aproximar una muestra de la distribución conjunta a posteriori. En este algoritmo se muestra sucesivamente las distribuciones condicionales, modificando las distribuciones en cada paso. Para más detalles ver el capítulo 10.*

## Algoritmo Gibbs

1.  $t = 0$ .
2. Elegir un valor inicial  $\mu^{(0)}$ .
3. Muestrear  $\phi^{(t+1)} \sim \mathcal{G}\left(\frac{a+n}{2}, \frac{b+(n-1)s^2+n(\mu^{(t)}-\bar{x})^2}{2}\right)$
4. Muestrear  $\mu^{(t+1)} \sim \mathcal{N}\left(\frac{\delta m+n\phi^{(t+1)}\bar{x}}{\delta+n\phi^{(t+1)}}, \frac{1}{\delta+n\phi^{(t+1)}}\right)$
5.  $t = t + 1$ .
6. Si  $t < T$ , ir a 2.

*Para  $T$  grande, los valores muestreados aproximan una muestra de la verdadera distribución a posteriori.*

**Ejemplo 29** *Se quiere comprobar si una hormona ayuda incrementar la producción promedio de trigo,  $\mu$  sobre su nivel estandar de 100 unidades por kilometro cuadrado.*

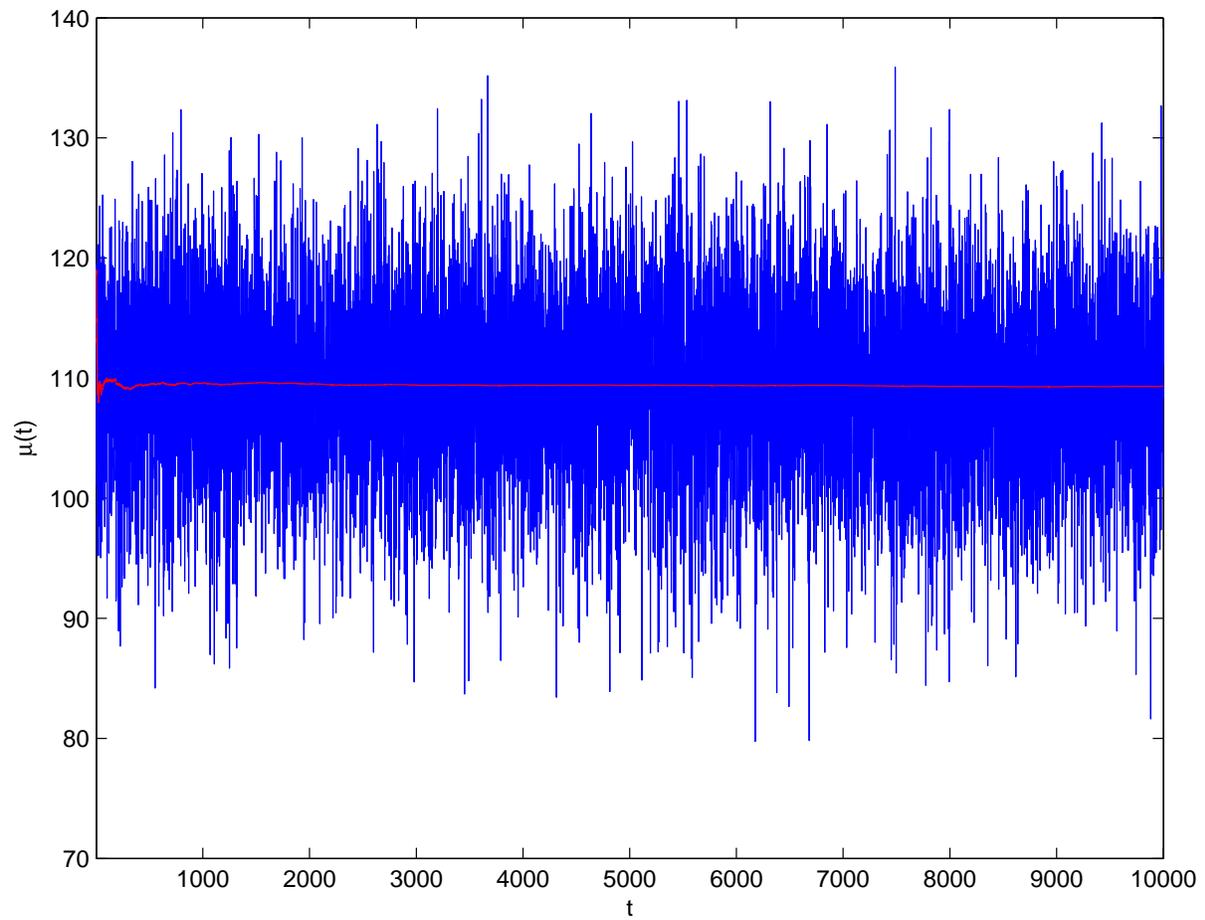
*Sea la distribución a priori  $\mu \sim \mathcal{N}(100, 100)$  con precisión  $\phi \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right)$ .*

*12 campos de  $1\text{km}^2$  tratadas con la hormona dan los siguientes cantidades de trigo:*

140, 103, 73, 171, 137, 91, 81, 157, 146, 69, 121, 134

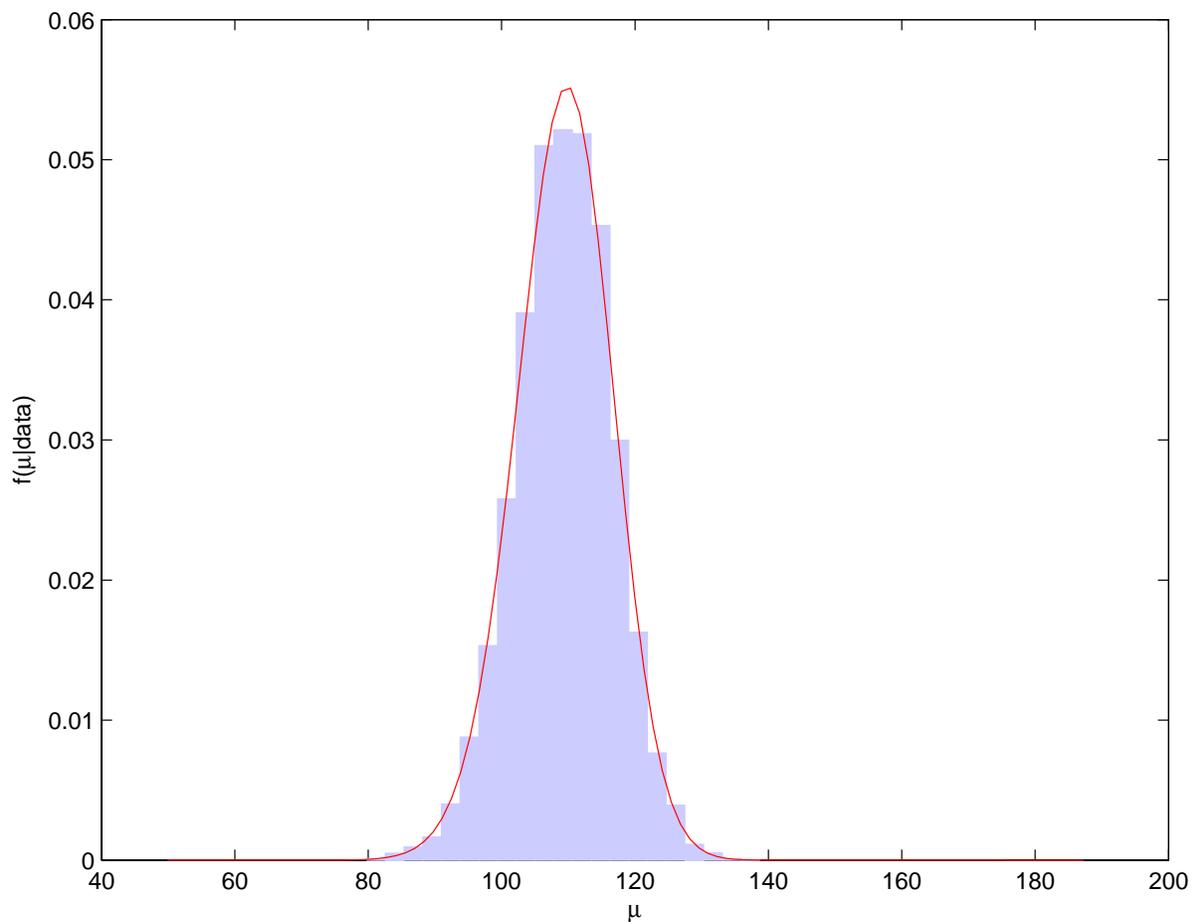
*Se quiere comprobar si hay suficiente evidencia a favor de la eficacia del tratamiento.*

*En la siguiente página se muestran los valores de  $\mu$  generados frente al número de iteraciones y la media estimada de  $\mu$  (en rojo) frente al número de iteraciones.*



*Se ve que la media estimada converge rápidamente.*

*El siguiente diagrama es un histograma de los valores generados y una estimación la densidad a posteriori de  $\mu$ .*



*Se estima la media a posteriori de  $\mu$  en 109,3 y la probabilidad de que  $\mu > 100$  en 0,895.*

## Inferencia para la distribución normal multivariante

Sea  $\mathbf{X}|\boldsymbol{\mu}, \mathbf{V} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{V})$  es decir que

$$f(\mathbf{x}|\boldsymbol{\mu}, \mathbf{V}) \propto |\mathbf{V}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$$

donde  $\mathbf{x} = (x_1, \dots, x_k)^T$ .

Dada una muestra  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ , se quiere hacer inferencia sobre  $\boldsymbol{\mu}, \mathbf{V}$ . La verosimilitud es

$$l(\boldsymbol{\mu}, \mathbf{V}|\mathbf{x}) \propto |\mathbf{V}|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})\right)$$

El análisis conjugada es parecida al análisis univariante. Sea

$$\begin{aligned}\boldsymbol{\mu}|\mathbf{V} &\sim \mathcal{N}\left(\mathbf{m}, \frac{1}{c}\mathbf{V}\right) \\ \mathbf{V} &\sim \mathcal{WI}(\alpha, \mathbf{W}^{-1})\end{aligned}$$

una distribución Wishart invertida, es decir que

$$f(\mathbf{V}) \propto |\mathbf{V}|^{-\frac{\alpha+k+1}{2}} \exp\left(-\frac{1}{2}\text{tr}(\mathbf{WV}^{-1})\right)$$

La densidad a priori conjunta se llama **la densidad normal Wishart invertida**

$$f(\boldsymbol{\mu}, \mathbf{V}) \propto |\mathbf{V}|^{-\frac{\alpha+k+2}{2}} \exp\left(-\frac{1}{2} \left[ \text{tr}(\mathbf{WV}^{-1}) + c(\boldsymbol{\mu} - \mathbf{m})^T \mathbf{V}^{-1} (\boldsymbol{\mu} - \mathbf{m}) \right]\right)$$

donde  $k = \dim(\boldsymbol{\mu})$ .

Dada esta distribución, la inferencia es conjugada, como se demuestra el siguiente teorema.

## Teorema 9

$$\begin{aligned}\boldsymbol{\mu}|\mathbf{V}, \mathbf{x} &\sim \mathcal{N}\left(\mathbf{m}^*, \frac{1}{c^*}\mathbf{V}\right) \\ \mathbf{V}|\mathbf{x} &\sim \mathcal{WI}\left(\alpha^*, \mathbf{W}^{*-1}\right)\end{aligned}$$

donde

$$\begin{aligned}\mathbf{m}^* &= \frac{c\mathbf{m} + n\bar{\mathbf{x}}}{c + n} \\ c^* &= c + n \\ \alpha^* &= \alpha + n \\ \mathbf{W}^* &= \mathbf{W} + \mathbf{S} + \frac{cn}{c + n}(\bar{\mathbf{x}} - \mathbf{m})(\bar{\mathbf{x}} - \mathbf{m})^T\end{aligned}$$

donde  $\mathbf{S} = \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$ .

La demostración del resultado es muy parecido al univariable, sustituyendo formas cuadráticas por cuadrados.

## Demostración

$$\begin{aligned}
 f(\boldsymbol{\mu}, \mathbf{V}|\mathbf{x}) &\propto |\mathbf{V}|^{-\frac{\alpha+k+2}{2}} \exp\left(-\frac{1}{2} [\text{tr}(\mathbf{WV}^{-1}) + \right. \\
 &\quad \left. c(\boldsymbol{\mu} - \mathbf{m})^T \mathbf{V}^{-1}(\boldsymbol{\mu} - \mathbf{m})] \right) \times \\
 &\quad |\mathbf{V}|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})^T \mathbf{V}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})\right) \\
 &\propto |\mathbf{V}|^{-\frac{\alpha+k+n+2}{2}} \exp\left(-\frac{1}{2} [\text{tr}(\mathbf{WV}^{-1}) + \right. \\
 &\quad (c+n)\boldsymbol{\mu}^T \mathbf{V}^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T \mathbf{V}^{-1}(c\mathbf{m} + n\bar{\mathbf{x}}) - \\
 &\quad \left. (c\mathbf{m} + n\bar{\mathbf{x}})^T \mathbf{V}^{-1}\boldsymbol{\mu} + c\mathbf{m}^T \mathbf{V}^{-1}\mathbf{m} + n^2\bar{\mathbf{x}}^T \mathbf{V}^{-1}\bar{\mathbf{x}} \right) \\
 &\propto |\mathbf{V}|^{-\frac{\alpha^*+k+2}{2}} \exp\left(-\frac{1}{2} [\text{tr}(\mathbf{WV}^{-1}) + c\mathbf{m}^T \mathbf{V}^{-1}\mathbf{m} + \right. \\
 &\quad \left. n^2\bar{\mathbf{x}}^T \mathbf{V}^{-1}\bar{\mathbf{x}} + c^*(\boldsymbol{\mu} - \mathbf{m}^*)^T \mathbf{V}^{-1}(\boldsymbol{\mu} - \mathbf{m}^*) - \right. \\
 &\quad \left. c^*\mathbf{m}^{*T} \mathbf{V}^{-1}\mathbf{m}^* \right] \right) \\
 &\propto |\mathbf{V}|^{-\frac{\alpha^*+k+2}{2}} \exp\left(-\frac{1}{2} [\text{tr}(\mathbf{WV}^{-1}) + \right. \\
 &\quad \frac{cn}{c+n}(\mathbf{m} - \bar{\mathbf{x}})^T \mathbf{V}^{-1}(\mathbf{m} - \bar{\mathbf{x}}) + \\
 &\quad \left. c^*(\boldsymbol{\mu} - \mathbf{m}^*)^T \mathbf{V}^{-1}(\boldsymbol{\mu} - \mathbf{m}^*)] \right)
 \end{aligned}$$

$$f(\boldsymbol{\mu}, \mathbf{V}|\mathbf{x}) \propto |\mathbf{V}|^{-\frac{\alpha^*+k+2}{2}} \exp\left(-\frac{1}{2} [\text{tr}(\mathbf{W}^*\mathbf{V}^{-1}) + c^*(\boldsymbol{\mu} - \mathbf{m}^*)^T \mathbf{V}^{-1}(\boldsymbol{\mu} - \mathbf{m}^*)]\right)$$

que es el núcleo de la distribución normal Wishart invertida.

◇

**Ejemplo 30 El modelo lineal normal** *Supongamos el modelo lineal*

$$\mathbf{Y}|\mathbf{X}, \boldsymbol{\theta} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\theta}, \mathbf{C})$$

*donde  $\mathbf{X}$  y  $\mathbf{C}$  son conocidos.*

*Tomando una distribución a priori  $\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{m}, \mathbf{V})$ , tenemos un modelo lineal de dos etapas (Lindley y Smith 1972).*

*Es fácil demostrar que dados los datos  $\mathbf{Y} = \mathbf{y}$ , tenemos la distribución final*

$$\boldsymbol{\theta}|\mathbf{x}, \mathbf{y} \sim \mathcal{N}(\mathbf{m}^*, \mathbf{V}^*)$$

*donde*

$$\begin{aligned}\mathbf{V}^* &= (\mathbf{V}^{-1} + \mathbf{X}^T \mathbf{C}^{-1} \mathbf{X})^{-1} \\ \mathbf{m}^* &= \mathbf{V}^* (\mathbf{V}^{-1} \mathbf{m} + \mathbf{X}^T \mathbf{C}^{-1} \mathbf{y})\end{aligned}$$

*Ver el Capítulo 7 para más detalles.*

## Aplicación 3: Inferencia clásica y bayesiana para la distribución media normal

Seguimos un trabajo de Wiper et al (2005).

### ¿Cómo es la distribución media normal?

Si  $Z \sim \mathcal{N}(0, 1)$ , luego  $|X|$  tiene una distribución media normal estandar  $X \sim \mathcal{MN}(0, 1)$ .

$$f(x) = 2\phi(x) \quad \text{para } x > 0, \text{ donde}$$
$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

## La distribución $\mathcal{MN}(\xi, \eta)$

Generalizando, si  $X = \xi + \eta|Z|$ , entonces se dice que  $X|\eta, \xi \sim \mathcal{HN}(\xi, \eta)$  y

$$f(x|\xi, \eta) = \frac{2}{\eta} \phi\left(\frac{x - \xi}{\eta}\right) \quad \text{para } x > \xi.$$

Se puede derivar la media de la distribución de las propiedades de la distribución normal truncada.

$$E[X|\xi, \eta] = \xi + \eta\sqrt{\frac{2}{\pi}}.$$

## ¿Porqué es interesante la distribución media normal?

Es un modelo para datos truncados. Se ha aplicado en

- distorsión de fibras (Haberle 1991).
- dispersión de moscas (Dobzhansky y Wright 1947).
- índices de grasa corporal de atletas (Pewsey 2002, 2004).
- fronteras estocásticas (Aigner et al 1977 y Meeusen y van den Broeck, 1977).

## Inferencia clásica

Dada una muestra  $\mathbf{x} = (x_1, \dots, x_n)$  se quiere estimar  $\xi$  (y  $\eta$ ). Ahora, Pewsey (2002,2004) demuestra que:

- El MLE de  $\xi$  es  $\hat{\xi} = x_{(1)}$ , es decir el mínimo de los datos.
- Se estima  $\eta$  mediante

$$\hat{\eta} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x_{(1)})^2}.$$

- Estas estimadores son sesgadas. Pewsey (2002,2004) proporciona estimadores corregidas:

$$\hat{\xi}_{BC} = x_{(1)} - \hat{\eta}_{BC} \Phi^{-1} \left( \frac{1}{2} + \frac{1}{2n} \right) \quad \text{donde}$$

$$\hat{\eta}_{BC} = \sqrt{\frac{n}{n-1}} \hat{\eta}$$

## Intervalos de confianza

- Un intervalo asintótico (no corregido) de  $100(1 - \alpha)\%$  confianza para  $\xi$ :

$$x_{(1)} + \log(\alpha)\hat{\eta}\Phi^{-1}\left(\frac{1}{2} + \frac{1}{2n}\right) < \xi < x_{(1)}$$

- Un intervalo corregido (de 2 lados) es

$$x_{(1)} + \log\left(\frac{\alpha}{2}\right)\hat{\eta}_{BC}\Phi^{-1}\left(\frac{1}{2} + \frac{1}{2n}\right) < \xi < \\ x_{(1)} + \log\left(1 - \frac{\alpha}{2}\right)\hat{\eta}_{BC}\Phi^{-1}\left(\frac{1}{2} + \frac{1}{2n}\right).$$

- Un intervalo de 1 lado

$$x_{(1)} + \log(\alpha)\hat{\eta}_{BC}\Phi^{-1}\left(\frac{1}{2} + \frac{1}{2n}\right) < \xi < x_{(1)}.$$

Pewsey (2002,2004) demuestra que sus estimadores tienen mejores propiedades que los EML pero ¿con inferencia bayesiana cómo va?

## Inferencia bayesiana para la distribución media normal

En primer lugar, se reparameteriza el modelo definiendo  $\tau = \frac{1}{\eta^2}$  para que

$$f(x|\xi, \tau) = \sqrt{\tau} \phi(\sqrt{\tau}(x - \xi)).$$

Ahora se puede definir una distribución a priori para  $\xi, \tau$  que generaliza la distribución normal gamma.

## La distribución normal gamma truncada a la derecha (RTNG)

$$\xi, \tau \sim \mathcal{RTNG}(\xi_0, m, \alpha, a, b) \quad \text{si}$$

$$f(\xi, \tau) = \frac{1}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(\alpha a)}}\right)} \frac{(b/2)^{a/2}}{\Gamma(a/2)} \sqrt{\frac{\alpha}{2\pi}} \tau^{\frac{a+1}{2}-1} \exp\left\{-\frac{\tau}{2} [b + \alpha(\xi - m)^2]\right\}$$

para  $\xi < \xi_0$  donde  $\Phi_d(\cdot)$  es la función de distribución de t de Student con  $d$  grados de libertad:

$$\Phi_d(z) = \int_{-\infty}^z \phi_d(y) dy \quad y$$

$$\phi_d(y) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \frac{1}{\sqrt{\pi d}} \left(1 + \frac{y^2}{d}\right)^{-\frac{d+1}{2}}$$

## ¿Porqué esta distribución?

Recordamos la inferencia conjugada para la normal. La distribución marginal de la media  $\mu$  es  $t$  y la distribución de  $\phi$  dado  $\mu$  es gamma. Para la distribución RTNG se tiene:

$$\begin{aligned} f(\xi) &= \frac{1}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(a\alpha)}}\right)} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)} \sqrt{\frac{\alpha}{b\pi}} \\ &\quad \left(1 + \frac{1}{a} \left(\frac{\xi - m}{\sqrt{b/(a\alpha)}}\right)^2\right)^{-\frac{a+1}{2}} \\ &= \sqrt{\frac{a\alpha}{b}} \frac{\phi_a\left(\frac{\xi - m}{\sqrt{b/(a\alpha)}}\right)}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(a\alpha)}}\right)} \quad \text{for } \xi < \xi_0 \\ \tau|\xi &\sim \mathcal{G}\left(\frac{a}{2}, \frac{b + \alpha(\xi - m)^2}{2}\right) \end{aligned}$$

**¿Cuál es la distribución marginal de  $\tau$ ?**

Es una distribución gamma modulada por una gaussiana  $\tau \sim \mathcal{GMG}(\sqrt{\alpha}(\xi_0 - m), a, b)$ , es decir

$$f(\tau) = \frac{\Phi(\sqrt{\alpha}\tau(\xi_0 - m)) \left(\frac{b}{2}\right)^{\frac{a}{2}}}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(a\alpha)}}\right) \Gamma\left(\frac{a}{2}\right)} \tau^{\frac{a}{2}-1} e^{-\frac{b}{2}\tau}.$$

Ambas distribuciones marginales parecen un poco chungas pero ...

## Tienen momentos

$$E[\xi] = m - \sqrt{\frac{b}{a\alpha}} \frac{a + \left(\frac{\xi_0 - m}{\sqrt{b/(a\alpha)}}\right)^2 \phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(a\alpha)}}\right)}{a - 1 \Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(a\alpha)}}\right)}$$

para  $a > 1$  y además se puede escribir la varianza de forma explícita.

## Momentos de $\tau$

$$E[\tau] = \frac{\Phi_{a+2}\left(\frac{\xi_0 - m}{\sqrt{b/(\alpha(a+2))}}\right) a}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(\alpha a)}}\right) b}$$

$$E[\tau^2] = \frac{\Phi_{a+4}\left(\frac{\xi_0 - m}{\sqrt{b/(\alpha(a+4))}}\right) a(a+2)}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(\alpha a)}}\right) b^2}$$

Es posible calcular todos los momentos positivos y negativos de  $\tau$  y luego de  $\eta$ .

## La distribución RTNG es conjugada con la distribución media normal

Si  $X|\xi, \tau \sim \mathcal{MN}\left(\xi, \frac{1}{\tau}\right)$  y  $\xi, \tau \sim \mathcal{RTNG}(\xi_0, m, \alpha, a, b)$  luego la distribución marginal de  $X$  es

$$f(x) = \frac{2\sqrt{\frac{a\alpha}{b(\alpha+1)}} \Phi_{a+1}\left(\frac{\min\{\xi_0, x\} - \frac{\alpha m + x}{\alpha + 1}}{\sqrt{\frac{b + \frac{\alpha}{\alpha+1}(x-m)^2}{(\alpha+1)(a+1)}}}\right)}{\Phi_a\left(\frac{\xi_0 - m}{\sqrt{b/(\alpha a)}}\right) \phi_a\left(\frac{x - m}{\sqrt{\frac{b(\alpha+1)}{\alpha a}}}\right)}$$

y se puede derivar una expresión explícita para su media.

Dada una muestra  $\mathbf{x} = x_1, \dots, x_n$ , la distribución a posteriori de  $\xi, \tau | \mathbf{x}$  es RTNG,

$$\xi, \tau | \mathbf{x} \sim \mathcal{RTNG}(\xi^*, m^*, \alpha^*, a^*, b^*)$$

donde

$$\begin{aligned}\xi^* &= \text{mín}\{x_{(1)}, \xi_0\} \\ m^* &= \frac{\alpha m + n \bar{x}}{\alpha + n} \\ \alpha^* &= \alpha + n \\ a^* &= a + n \\ b^* &= b + (n - 1)s^2 + \frac{\alpha n}{\alpha + n}(m - \bar{x})^2.\end{aligned}$$

## Resultados bajo una distribución a priori no informativa

Siendo

$$f(\xi, \tau) \propto \frac{1}{\tau}$$

entonces

$$\xi, \tau | \mathbf{x} \sim \mathcal{RTNG} \left( x_{(1)}, \bar{x}, n, n - 1, (n - 1)s^2 \right).$$

Desde ahora se va a suponer esta distribución a posteriori con el objetivo de comparar los estimadores bayesianos con los clásicos.

## Propiedades clásicas de estimadores bayesianas

Se simularon 1000 muestras de varios tamaños  $n$  de  $\mathcal{MN}(0, 1)$ . La tabla compara el sesgo esperado y la raíz del error medio cuadrático de las estimadores bayesianas y clásicas de  $\xi$ .

$n$	$X_{(1)}$		$\hat{\xi}_{BC}$		$E[\xi \mathbf{X}]$	
	bias	RMSE	bias	RMSE	bias	RMSE
5	0,2156	0,2882	-0,0031	0,2229	-0,0628	0,2440
10	0,1153	0,1575	-0,0015	0,1157	-0,0111	0,1192
20	0,0591	0,0823	-0,0011	0,0591	-0,0029	0,0597
50	0,0246	0,0345	-0,0001	0,0245	-0,0005	0,0245
100	0,0125	0,0176	$3 \times 10^{-5}$	0,0125	-0,0002	0,0125
1000	0,0012	0,0018	$9 \times 10^{-7}$	0,0012	$2 \times 10^{-6}$	0,0012

Los resultados son parecidos al estimador de Pewsey. La media a posteriori de  $\eta$  también tiene poco sesgo.

## ¿Y la cobertura?

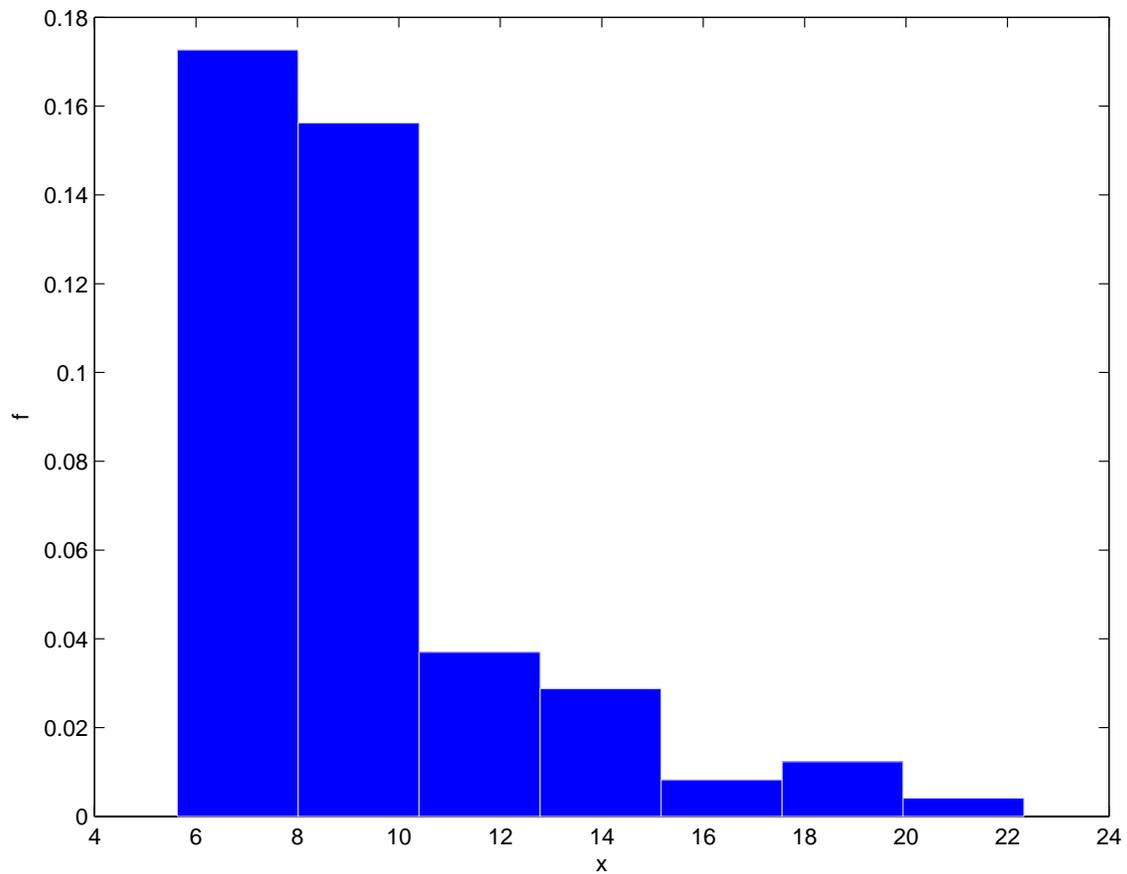
Comparamos los dos intervalos propuestos por Pewsey (2002,2004) con el intervalo bayesiano de máxima densidad a posteriori para  $\xi$ .

$n$	Frequentist intervals				Bayesian interval	
	2 sided		1 sided		coverage	length
	coverage	length	coverage	length	coverage	length
5	0,911	0,801	0,902	0,655	0,951	0,910
10	0,935	0,428	0,931	0,349	0,950	0,398
20	0,944	0,222	0,943	0,181	0,951	0,191
50	0,948	0,091	0,947	0,074	0,950	0,076
100	0,948	0,046	0,947	0,037	0,950	0,038
1000	0,951	0,005	0,951	0,004	0,949	0,004

Los intervalos bayesianos tienen buenas propiedades.

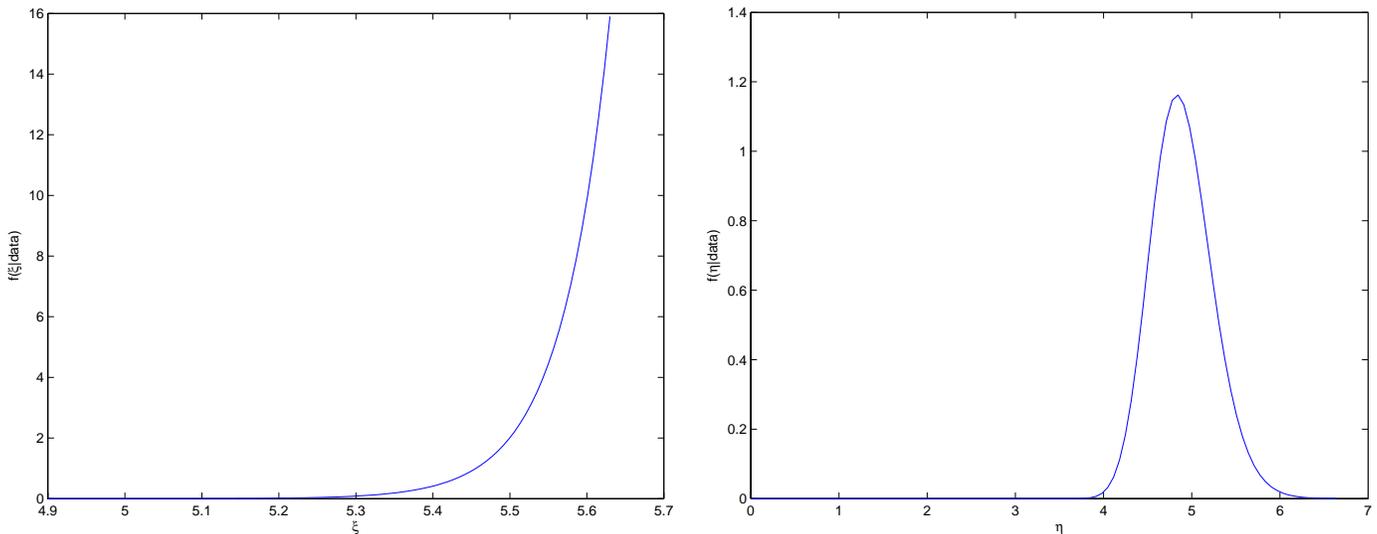
## Un ejemplo con datos reales

Datos de índices de grasa corporal de atletas.



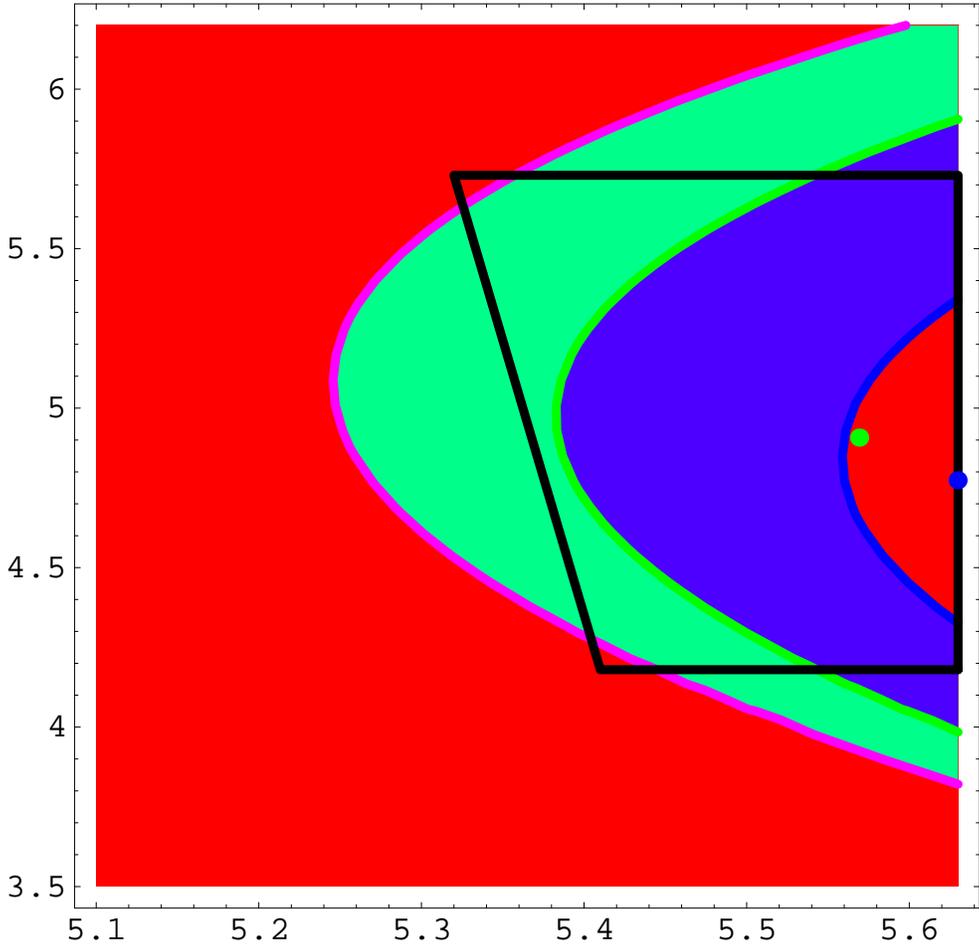
Parece razonable modelizar estos datos con una distribución media normal.

## Densidades a posteriori de $\xi$ y $\eta$



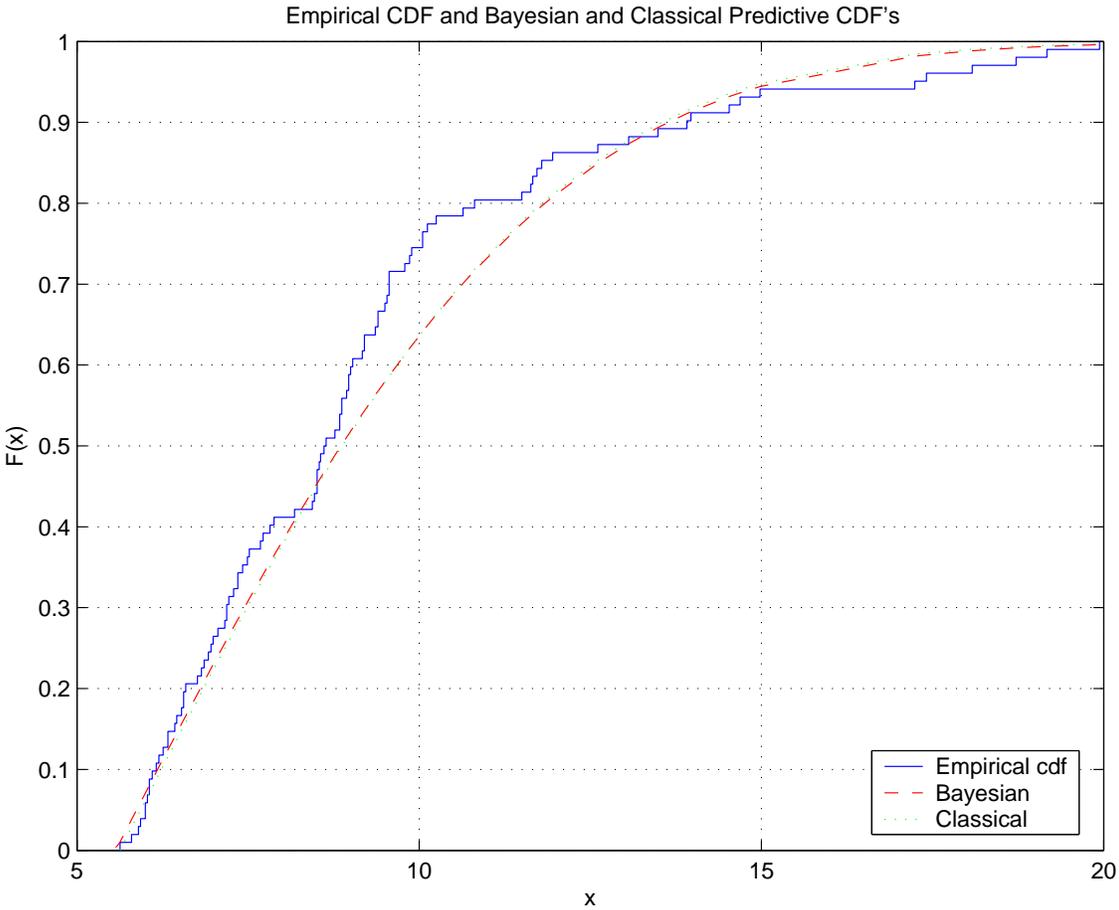
Las medias a posteriori ( $E[\xi|\mathbf{x}] = 5,57$ ) son iguales a las estimadores de Pewsey. Los intervalos de confianza ( $5,41 < \xi < 5,63$ ) y de credibilidad ( $5,44 < \xi < 5,63$ ) son un poco distintos.

# Regiones de confianza y de credibilidad



Son más distintos aún.

# La distribución predictiva



Los ajustes del modelo clásico y del modelo bayesiano son parecidos pero no tan buenos.

## Extensiones

Ver Wiper et al (2005).

- Considerar el modelo medio t.
- ¿Cómo elegir el modelo?
- Distribuciones elípticas.