MAT703. Parcial 2 Fausto Fabian Crespo Fernandez

Fausto Fabian Crespo Fernandez

junio 2016

0.1. Ejercicio 1

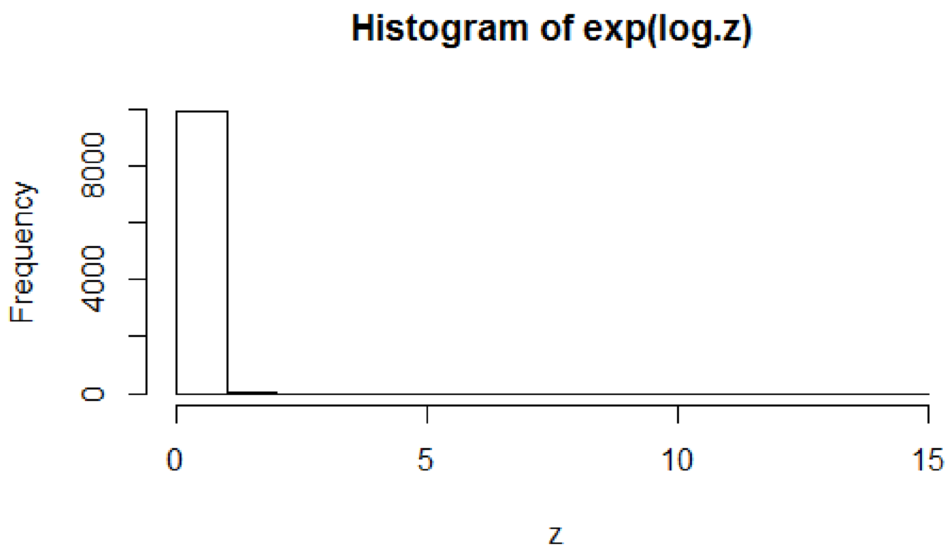
Solución Tenemos que $X \sim N(8, 14)$ y $Y \sim N(-3, 50)$ entonces definimos variable aleatoria $Z = \frac{X+1}{X^2+Y^4}$ y $W = X$ queremos hallar la densidad de esta variable aleatoria Z.
podemos encontrar primero la densidad conjunta de Z, W y luego integrar por $W = X$.Tenemos $p(x, y) = p(z, w)|det \begin{pmatrix} \frac{dZ}{dX} & \frac{dZ}{dY} \\ \frac{dW}{dX} & \frac{dW}{dY} \end{pmatrix}|$
$= p(z, w)|det \begin{pmatrix} \frac{x^2+y^4-2x(x+1)}{(x^2+y^4)^2} & \frac{-4y^3(x+1)}{(x^2+y^4)^2} \\ 1 & 0 \end{pmatrix}|$
$= p(z, w)\frac{|4y^3(x+1)|}{(x^2+y^4)^2}$ de donde
$p(z, w) = p(x, y)\frac{(x^2+y^4)^2}{|4y^3(x+1)|}$
y asumiendo que X y Y son independientes tenemos
$p(z, w) = \frac{1}{\sqrt{(2\pi*14)}}e^{\frac{-1}{2*14}(x-8)^2}\frac{1}{\sqrt{(2\pi*50)}}e^{\frac{-1}{2*50}(y+3)^2}\frac{(x^2+y^4)^2}{|4y^3(x+1)|}$ Ahora la densidad de Z es
$p(z) = \int_{-\infty}^{\infty} p(z, w)dw$
$p(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{(2\pi*14)}}e^{\frac{-1}{28}(x-8)^2}\frac{1}{\sqrt{(2\pi*50)}}e^{\frac{-1}{100}(y+3)^2}\frac{(x^2+y^4)^2}{|4y^3(x+1)|}dx$
$p(z) = \frac{1}{8\pi\sqrt{(50*14)}}e^{\frac{-1}{100}(y+3)^2}\int_{-\infty}^{\infty} e^{\frac{-1}{28}(x-8)^2}\frac{(x^2+y^4)^2}{|y^3(x+1)|}dx$ La media y la varianza de Z se pueden obtener como
$E(z) = \int_{-\infty}^{\infty} zp(z)dz$
$var(z) = E(z^2) - (E(z))^2$
También se puede hacer con
$E(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z(x, y)p(xy)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{x+1}{x^2+y^4}p(x)p(y)dxdy$

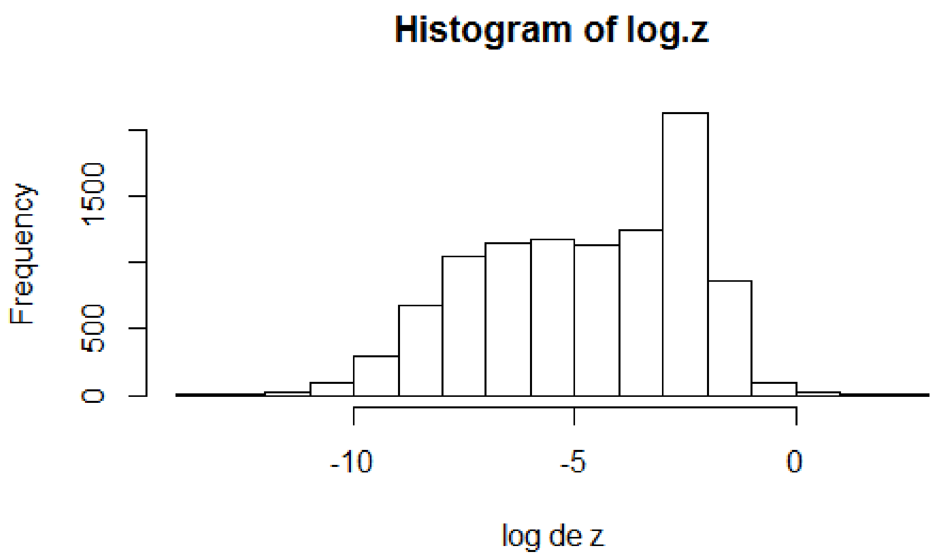
Otra forma es usar el polinomio de Taylor para aproximar el valor esperado y la varianza de Z. Si $T = (X, Y)$ y $\theta = E(T) = (8, -3)$ y tenemos la función $g(T) = \frac{X+1}{X^2+Y^4)}$ derivable
$g(T) \approx g(\theta) + \frac{\partial g}{\partial X}(\theta)(T_1 - \theta_1) + \frac{\partial g}{\partial Y}(\theta)(T_2 - \theta_2)$ de aqui sale que
$E(g(T)) \approx g(\theta) + \frac{\partial g}{\partial X}(\theta)E(T_1 - \theta_1) + \frac{\partial g}{\partial Y}(\theta)E(T_2 - \theta_2) = g(\theta)$
O sea $E(Z) \approx (8 + 1)/(8^2 + (-3)^4) = 0{,}06206897$. También la varianza es
$var(g(T)) \approx E((g(T) - g(\theta))^2) \approx E(\frac{\partial g}{\partial X}(\theta)(T_1 - \theta_1) + \frac{\partial g}{\partial Y}(\theta)(T_2 - \theta_2))$
$= (\frac{\partial g}{\partial X}(\theta))^2 var(T_1) + (\frac{\partial g}{\partial Y}(\theta))^2 var(T_2) + 2\frac{\partial g}{\partial X}(\theta)\frac{\partial g}{\partial Y}(\theta)Cov(T_1, T_2)$
$\approx 14 * \frac{(3^4-8^2+2*8)^2}{(8^2+3^4)^4} + 50 * \frac{(4*3^3*9)^2}{(8^2+3^4)^4} = 0{,}1068638$
El error que se comete con la aproximación de $g(T)$ es

$O[(\frac{1}{2}(T_i-\theta_i)^T H(\theta+c(T_i-\theta_i))(T_1-\theta_1))]$
con H la matriz Hessiana(matriz de segundas derivadas de z) y $c\in[0,1]$.
Tenemos :
$H(\theta)=\begin{pmatrix}-2642/3048625 & -15444/3048625\\ -15444/3048625 & 69012/3048625\end{pmatrix}=\begin{pmatrix}-0{,}0009 & -0{,}0051\\ -0{,}0051 & 0{,}0226\end{pmatrix}$

Podemos hacer una simulacion con 100000 extracciones de X y Y y calcular Z, pero antes podemos definir $U=lnz=ln(x+1)+ln(x^2+y^4)$ esto es para evitar overflow o underflow. Esto se puede hacer porque el signo de Z es igual al signo de $x+1$(pues el denominador siempre es positivo) pero como el intervalo de confianza al 95 % de X es $[8-1{,}96*\sqrt{(}14),8+1{,}96*\sqrt{(}14)]=[qnorm(0{,}025,8,sqrt(14)),qnorm(0{,}975,8,sqrt(14))]=[0{,}6664863,15{,}33351]$ la mayoria de las veces los valores de X simulados serán positivos.Para Y el intervalo al 95 % es $[-16{,}85904,10{,}85904]$ y hay que tomar en cuenta que el denominador de Z puede ser pequenno y por eso es razonable tomar logaritmo de Z
Tenemos en R:
$n.sim=100000$
$x=rnorm(n.sim,8,sqrt(14))$
$y=rnorm(n.sim,-3,sqrt(50))$
$z=(x+1)/(x^2+y^4)$
$print(mean(z))$
$print(var(z))$
$log.z=log(x+1)-log(x^2+y^4)$
$hist(log.z,xlab="logdez",cex=2)$
$print(mean(log.z))$
$print(median(log.z))$
$deciles=quantile(log.z,probs=c(0{,}1,0{,}2,0{,}3,0{,}4,0{,}5,0{,}6,0{,}7,0{,}8,0{,}9))$
$print(deciles)$
$lines(density(log.z),col="red")$
$hist(exp(log.z),xlab="z",cex=2)$
$print(mean(exp(log.z)))$
$print(median(exp(log.z)))$
$deciles=quantile(exp(log.z),probs=c(0{,}1,0{,}2,0{,}3,0{,}4,0{,}5,0{,}6,0{,}7,0{,}8,0{,}9))$
$print(deciles)$
$lines(density(exp(log.z)),col="red")$

Con lo que se obtiene una media de z de 0.08357222 y varianza de z de 21.78559. Los gráficos

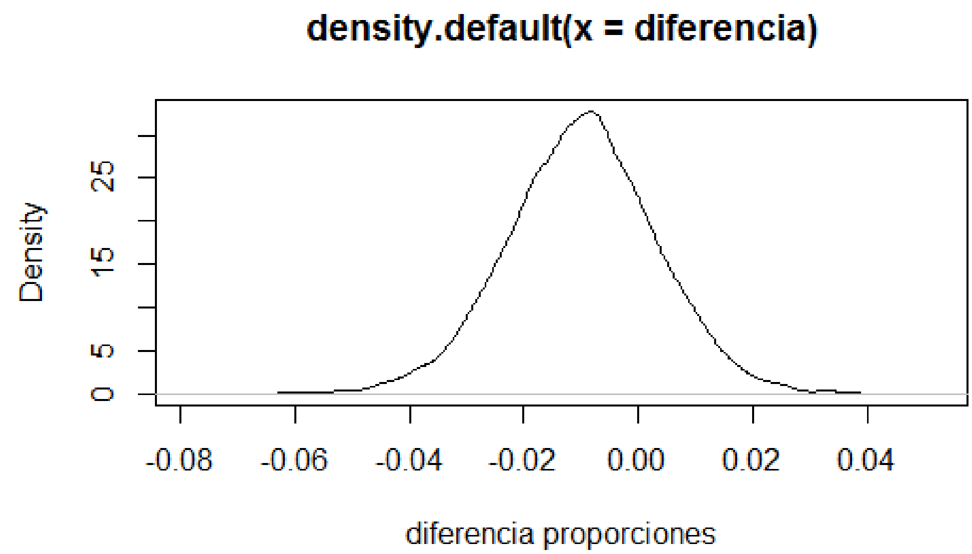

Histogram of exp(log.z)
Frequency
0
4000
8000
0
5
10
15
z


Histogram of log.z
Frequency
0
500
1500
-10
-5
0
log de z

0.2. Ejercicio 2

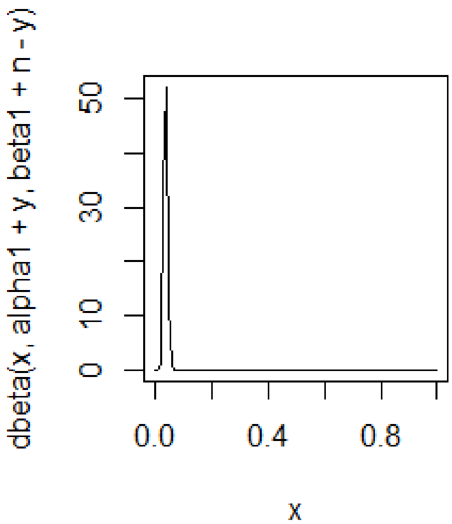
Solución

Tenemos una prueba de hipótesis $H_0 : p_1 < p_2$ donde p_1 y p_2 las probabilidades de muerte en el los dos hospitales. Tenemos que las muertes $y_i \sim Bin(p_i, n_i)$ o sea $y_1 \sim Bin(p_1, 430)$ y $y_2 \sim Bin(p_2, 390)$ los estimados para $\hat{p_1} = 12/430 = 0{,}0279$ y $\hat{p_2} = 12/430 = 0{,}041$.
Una primera opción de analisis es considerar que p_1 y p_2 tienen a prioris $Beta(\alpha_i, \beta_i)$(por ejemplo no informativas $Beta(1, 1)$ cada uno)y la a posteriori es un producto de dos dsitribuciones betas independientes $Beta(\alpha_i + y_i, \beta_i + y_)$ y analizamos la diferencia de p_1 y p_2 con 10000 simulaciones. En R:
$y1 < -12;$
$n1 < -430;$
$alpha1 < -1;$
$beta1 < -1$
$y2 < -16; n2 < -390; alpha2 < -1; beta2 < -1$
$p1 < -rbeta(10000, y1 + alpha1, n - y1 + beta1)$
$p2 < -rbeta(10000, y2 + alpha2, n - y2 + beta2)$
$diferencia < -p1 - p2$
$plot(density(diferencia), xlab = "diferenciaproporciones")$
$quantile(diferencia, c(,025, ,975))$
$mean(diferencia)$
$median(diferencia)$
$print(sum(diferencia < 0))$
$print(mean(diferencia > 0))$
Con lo que se tiene



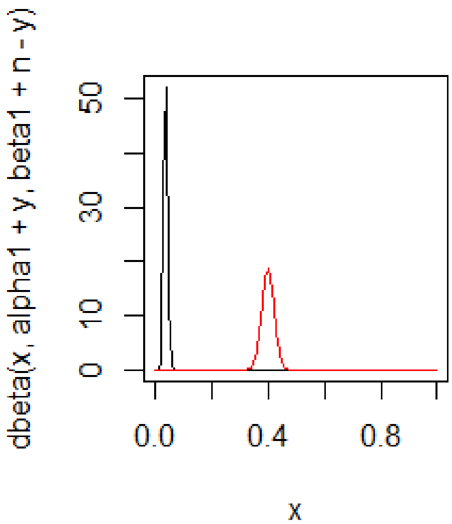
y la media de la diferencia es -0.009844688 con intervalo de confianza :
2.5-0.03606110 0.01562444 y la media de los valores negativos de la diferencia es 0.7831 y la media de los valores positivos de la diferencia es 0.2169 que es mayor de 0.05, por tanto según esto no hay evidencia para aceptar H_0 y se rechaza.
Otro enfoque es un modelo jerárquico:podemos asumir que la a priori común para los parámetros p_1, p_2 es una $Beta(\alpha, \beta)$ y podemos estimar los α, β a partir de los datos que tenemos ya que tenemos la media y la varianza de los datos e igualamos con la media y la varianza de una distribución $Beta$
En R
$y = c(12, 16)$
$n = c(430, 390)$
$proporciones = y/n$
$print(proporciones)$
$esperado.p = mean(proporciones)$
$varianza.p = var(proporciones)$
$normaDiferencia < -function(vectorparametros)$
{
$alpha = vectorparametros[1]$
$beta = vectorparametros[2]$
$denominador = alpha + beta$

$return((alpha/denominador - esperado.p)^2 + (alpha*beta/(denominador^2*(denominador+1)) - varianza.p)^2)$
$\}$
$library(optimx)$
$optimos = optimx(par = c(0{,}5, 0{,}2), normaDiferencia)$
$print(optimos)$
$alpha1 = optimos[1, 1]$
$beta1 = optimos[1, 2]$
$print(alpha1)$
$print(beta1)$
Con lo que se obtiene $(\alpha, \beta) = (12{,}95242, 362{,}8079)$. con esto podemos hacer una estimación a posteriori para p_2
En R
$alpha2 = alpha1 + 16$
$beta2 = beta1 + 390 - 16$
$p2.estimado = alpha2/(alpha2 + beta2)$
$print(p2.estimado)$
lo que da 0,03780873 La densidad a posteriori y un intervalo de confianza al 95 % para p es en R:
$y = 16$
$n = 390$
$par(pty = "s")$
$curve(dbeta(x, alpha1 + y, beta1 + n - y), from = 0, to = 1)$
$intervalo1 = c(qbeta(0{,}015, alpha1 + y, beta1 + n - y), qbeta(0{,}985, alpha1 + y, beta1 + n - y))$
$print(intervalo1)$
lo que da



y el intervalo de confianza es [0,024386210,05420762]
La moda de la a posteriori es 0.4 con el código en R:
$beta2 < -function(x)\{$
$return(-dbeta(x, alpha1 + y, beta1 + n - y))$
$\}$
$optimos1 = optimx(par = 0{,}4, beta2)$
$moda = optimos1\$p1[1]$
$print(moda)$

Podemos hacer una aproximacion de la distribución a posteriori a la normal y calcular el intervalo de confianza al 95 % para esa aproximación(media la moda y varianza el inverso de la información de Fisher). En R:
$informacion = (alpha1 + y - 1)/moda^2 + (beta1 + n - y - 1)/(1 - moda)^2$
$varianza = 1/informacion$
$curve(dnorm(x, moda, sqrt(varianza)), from = 0, to = 1, add = TRUE, col =$
$"red")$
$intervalo2 = c(qnorm(0{,}015, moda, sqrt(varianza)), qnorm(0{,}985, moda, sqrt(varianza)))$
$print(intervalo2)$



y el intervalo es [0,3539280,446072].
Este último intervalo muy lejos del primer intervalo pues solo tenemos datos de dos hospitales.El primer intervalo de confianza [0,024386210,05420762] si incluye el valor 16/390=0.04102564 por lo tanto rechazamos la hipótesis nula H_0.

0.3. Ejercicio 3

Solución
Tenemos que $y_i|\nu_i \sim N(v_i, S-i^2)$ y que $\nu_i \sim N(\mu_1, \tau_1^2)$ para los controles $i = 1, 2, 4, 5, 6, 7$ y $v_i \sim N(\mu_2, \tau_2^2)$ para lo cohortes $i = 3, 8$, además sabemos que las a prioris son $\mu_1 \sim N(0, 10)$ y que $\mu_1 - \mu_2 \sim N(0, 0{,}1)$.Tenemos que los hiperparametros son $\mu_1, \mu_2, \tau_1, \tau 2$ que influuen en los parámetros ν_1, ν_2 o sea un modelo jerárquico.Además sabemos que $\frac{s_0}{(s_0+\tau_1)^2} \sim Pareto(\alpha, x_m)$ y $\frac{s_0}{(s_0+\tau_2)^2} \sim Pareto(\alpha, x_m)$. Tenemos que la a a priori conjunta es
$p(\nu_1, \ldots, \nu_8|\mu_1, \mu_2, \tau_1, \tau_2) = \prod_{i=1}^{8} N(\nu_i|\mu_e, \tau_e^2)$
$= N(\nu_3|\mu_2, \tau_2^2)N(\nu_8|\mu_2, \tau_2^2) \prod_{i=1, i\neq 3,8}^{8} N(\nu_i|\mu_1, \tau_1^2)$
y $p(\nu_1, \ldots, \nu_8)$
$= \int N(\nu_3|\mu_2, \tau_2^2)N(\nu_8|\mu_2, \tau_2^2) \prod_{i=1, i\neq 3,8}^{8} N(\nu_i|\mu_1, \tau_1^2) p(\mu_1, \mu_2, \tau_1, \tau_2) d(\mu_1, \mu_2, \tau_1, \tau_2)$
Tenemos que

$p(\nu_1,\ldots,\nu_8,\mu_1,\mu_2,\tau_1,\tau_2|y) \propto p(\mu_1,\mu_2,\tau_1,\tau_2)p(\nu_1,\ldots,\nu_8|\mu_1,\mu_2,\tau_1,\tau_2)p(y|\nu_1,\ldots,\nu_8)$
$\propto p(\mu_1,\mu_2,\tau_1,\tau_2)N(\nu_3|\mu_2,\tau_2^2)N(\nu_8|\mu_2,\tau_2^2)\prod_{i=1,i\neq 3,8}^{8} N(\nu_i|\mu_1,\tau_1^2)\prod_{i=1}^{8} N(y_i|\nu_i,s_i^2)$

y llamando a $x = (\nu_3-\mu_2,\nu_8-\mu_2)_T$ se puede agrupar términos y completando los cuadrados en el exponente para $\nu_i, i = 1,2,4,5,6,7$ y escribir
$p(\nu_1,\ldots,\nu_8,\mu_1,\mu_2,\tau_1,\tau_2|y) \propto$
$\prod_{i=1}^{8} N(\hat{\nu_i},V_i)$
$\hat{\nu_i} = \frac{\frac{1}{s_i^2}y_i+\frac{1}{\tau_1^2}\mu_1}{\frac{1}{s_i^2}+\frac{1}{\tau_1^2}}$ y $V_i^{-1} = \frac{1}{\frac{1}{s_i^2}+\frac{1}{\tau_1^2}}$ con $i = 1,2,4,5,6,7$
$\hat{\nu_i} = \frac{\frac{1}{s_i^2}y_i+\frac{1}{\tau_2^2}\mu_2}{\frac{1}{s_i^2}+\frac{1}{\tau_2^2}}$ y $V_i^{-1} = \frac{1}{\frac{1}{s_i^2}+\frac{1}{\tau_2^2}}$ con $i = 3,8$
por tanto tenemos producto de distribuciones independientes y
$\nu_i|\mu_1,\mu_2,\tau_1,\tau_2,y \sim N(\hat{\nu_i},V_i)$ con $i = 1,\ldots,8$
Ahora aplicando la propiedad de la distribución normal siguiente, tenemos que $y_i|\mu_1,\mu_2,\tau_1,\tau_2 \sim N(\nu_i,s_i^2+\tau_1^2)$ para $i = 1,2,4,5,6,7$ y
$y_i|\mu_1,\mu_2,\tau_1,\tau_2 \sim N(\nu_i,s_i^2+\tau_2^2)$ para $i = 3,8$
y $p(\mu_1,\mu_2,\tau_1,\tau_2|y) \propto N(\nu_1,s_1^2+\tau_2^2)N(\nu_2,s_2^2+\tau_2^2)\prod_{i=1,i\neq 3,8}^{8} N(\nu_i,s_i^2+\tau_1^2)$
y como $p(\mu_1,\mu_2,\tau_1,\tau_2|y) \propto p(\mu_1,\mu_2|\tau_1,\tau_2,y)p(\tau_1,\tau_2|y)$ si se asume una una distribución condicional a priori uniforme $p(\mu_1,\mu_2|\tau_1,\tau_2)$ entonces el logaritmo de la posterior es la suma de dos funciones cuadráticas indepenedientes en μ_1 y μ_2 respectivamente $p(\mu_1,\mu_2|\tau_1,\tau_2,y)$ es el producto de dos normales y
$\mu_1|\tau_1,\tau_2,y \sim N(\tilde{\mu_1},V_{\mu_1})$ y
$\mu_2|\tau_1,\tau_2,y \sim N(\tilde{\mu_2},V_{\mu_2})$
con $\tilde{\mu_1} = \frac{\sum_{i=1,i\neq 3,8}^{8}\frac{1}{\tau_1^2+s_i^2}y_i}{\sum_{i=1,i\neq 3,8}^{8}\frac{1}{\tau_1^2+s_i^2}}$ y $V_{\mu_1} = \sum_{i=1,i\neq 3,8}^{8}\frac{1}{\tau_1^2+s_i^2}$ y
$\tilde{\mu_2} = \frac{\sum_{i=1,i=3,8}^{8}\frac{1}{\tau_1^2+s_i^2}y_i}{\sum_{i=1,i=3,8}^{8}\frac{1}{\tau_2^2+s_i^2}}$ y $V_{\mu_2} = \sum_{i=1,i=3,8}^{8}\frac{1}{\tau_2^2+s_i^2}$
O sea que las distribuciones a posteriori de los hiperparámetros μ_1 y μ_2 condicionadas a τ_1,τ_2 y los datos y son normales .
La a posteriori para τ_1,τ_2 es
$p(\tau_1,\tau_2|y) = \frac{p(\mu_1,\mu_2,\tau_1,\tau_2|y)}{p(\mu_1,\mu_2|\tau_1,\tau_2,y)}$
$p(\tau_1,\tau_2|y) \propto (p(\tau_1)V_{\mu_1}^{1/2})(\prod_{i=1,i\neq 3,8}^{8}(s_i^2+\tau_1^2)^{-1/2}exp(-\frac{(y_i-\mu_1)^2}{2(s_i^2+\tau_1^2)}))(p(\tau_2)V_{\mu_2}^{1/2})(\prod_{i=1,i=3,8}^{8}(s_i^2+\tau_2^2)^{-1/2}exp(-\frac{(y_i-\mu_2)^2}{2(s_i^2+\tau_2^2)}))$ y las a prioris para τ_1 y τ_2 se pueden calcular con

$$p(\tau_1,\tau_2) = p(\frac{s_{01}}{(s_{01}+\tau_1)^2})p(\frac{s_{02}}{(s_{0a}+\tau_2)^2})|det\begin{pmatrix}\frac{-2}{(s_{01}+\tau_1)^3} & 0\\ 0 & \frac{-2}{(s_{02}+\tau_2)^3}\end{pmatrix}|$$

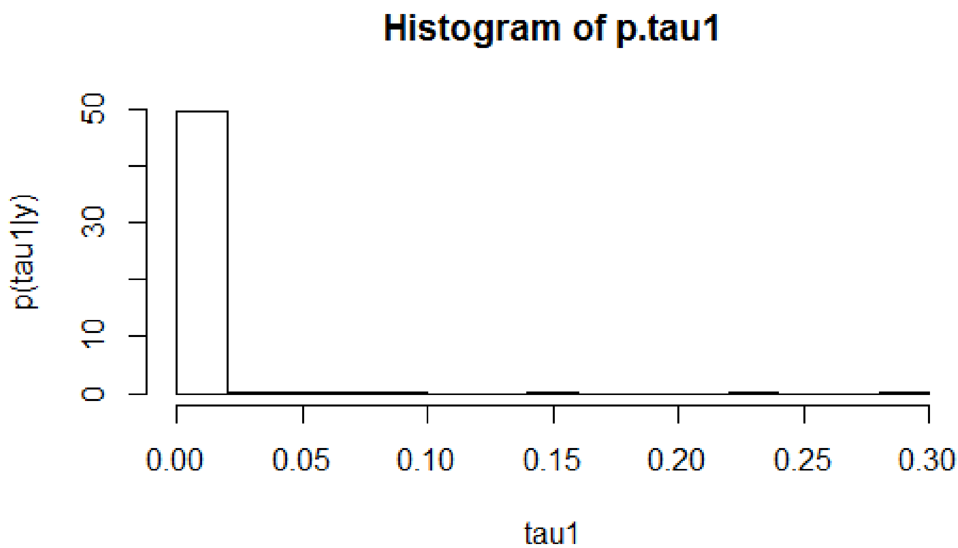
$= \frac{\alpha x_m^\alpha}{\tau_1^{\alpha+1}} \frac{\alpha x_m^\alpha}{\tau_1^{\alpha+1}} * 4 \frac{s_{01} s_{02}}{(s_{01}+\tau_1)^3 (s_{02}+\tau_2)^3}$

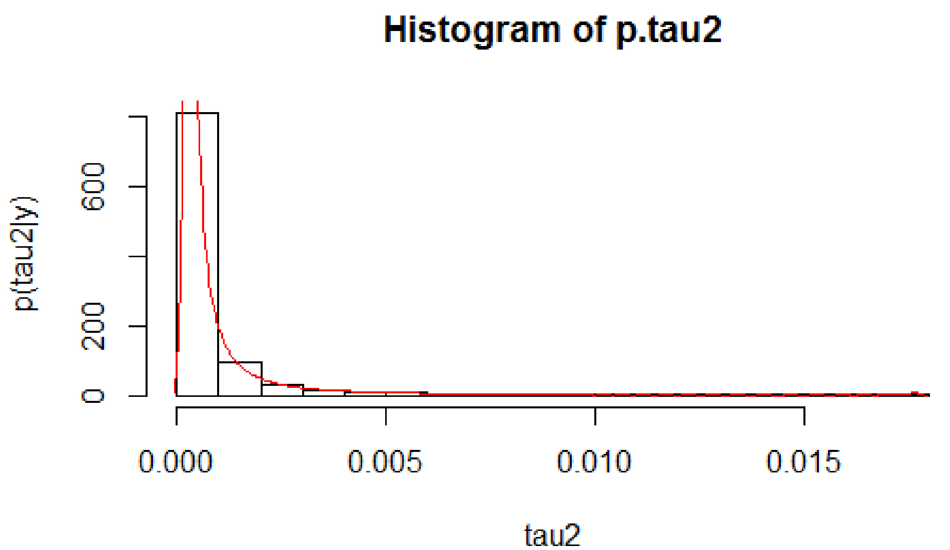
La a posteriori para μ_1, μ_2 es
$p(\mu_1, \mu_2|y) = \frac{p(\mu_1,\mu_2,\tau_1,\tau_2|y)}{p(\tau_1,\tau_2|\mu_1,\mu_2,y)}$. O sea tendríamos que hallar primero la distribucion condicional a posteriori $p(\tau_1, \tau_2|\mu_1, \mu_2, y)$.Pero tambien tenemos
$p(\mu_1, \mu_2|y) = p(\mu_1, \mu_2|\tau_1, \tau_2, y)p()$
Podemos hacer una simulación en R:
$y = c(0{,}405, 0{,}386, 0{,}698, 0{,}637, 0{,}247, 0{,}239, 0{,}148, 0{,}693, 0{,}236, 0{,}315, 0{,}278)$
$y.cohortes = y[c(3, 8)]$
$y.controles = y[c(1, 2, 4, 5, 6, 7)]$
$sd.y = c(0{,}695, 0{,}451, 0{,}730, 0{,}481, 0{,}134,$
$0{,}206, 0{,}163, 0{,}544, 0{,}246, 0{,}591, 0{,}487)$
$var.y = (sd.y)^2$
$sd.y.cohortes = sd.y[c(3, 8)]$
$sd.y.controles = sd.y[c(1, 2, 4, 5, 6, 7)]$
$var.y.cohortes = (sd.y.cohortes)^2$
$var.y.controles = (sd.y.controles)^2$
$s0.controles1 = 1/mean(1/var.y.controles)$
$s0.cohortes = 1/mean(1/var.y.cohortes)$
$J = length(y)$
$ycombinado = sum((1/var.y^2) * y)/sum(1/var.y^2)$
$print(ycombinado)$
$ycombinado.controles = sum((1/var.y.controles^2)*y.controles)/sum(1/var.y.controles^2)$
$print(ycombinado.controles)$
$ycombinado.cohortes = sum((1/var.y.cohortes^2)*y.cohortes)/sum(1/var.y.cohortes^2)$
$print(ycombinado.cohortes)$
$var.posteriori = 1/sum(1/var.y^2)$
$var.posteriori.controles = 1/sum(1/var.y.controles^2)$
$var.posteriori.cohortes = 1/sum(1/var.y.cohortes^2)$
$intervalo = c(ycombinado - 1{,}96 * sqrt(var.posteriori), ycombinado + 1{,}96 *$
$sqrt(var.posteriori))$
$print(var.posteriori)$
$print(intervalo)$
$intervalo.controles = c(ycombinado.controles - 1{,}96 * sqrt(var.posteriori.controles),$
$ycombinado.controles + 1{,}96 * sqrt(var.posteriori.controles))$
$print(var.posteriori.controles)$
$print(intervalo.controles)$

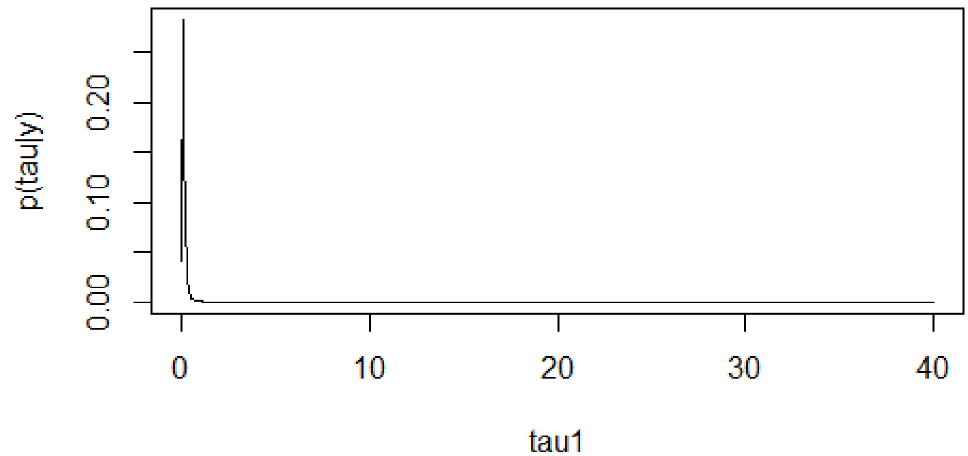
$intervalo.cohortes = c(ycombinado.cohortes - 1{,}96 * sqrt(var.posteriori.cohortes),$
$ycombinado.cohortes + 1{,}96 * sqrt(var.posteriori.cohortes))$
$print(var.posteriori.cohortes)$
$print(intervalo.cohortes)$
Primero hallamos los efectos combinados(ponderados por 1/varianza) y los efectos combinados en controles y cohortes. El $\bar{y}_{..}$ combinado total es 0.2233408, el de los controles es 0.2209555 y el de los cohortes es 0.6941785.Los intervalos de confianza al 95 % para el efecto combinado en los controles es $[0{,}1935653, 0{,}2483457]$ y para los cohortes es [0,18708881,2012682] por lo que se ve que hay solapamiento en los intervalos.Si hacemos un test de la estadística clásica para comprobar si todos los ν_j estiman la misma cantidad, hacemos un estadística χ_2 con 11-1 grados de libertad y estadístico $\sum_{j=1}^{11}(y_j - \bar{y}_{..})^2/s_j^2 = 17{,}36874$ con probabilidad de 0.9334073.Lo que indicaría que no rechazamos la hipótesis nula, o sea que si están estimando los mismos valores. En R
$estadistico.chi = sum((y - ycombinado)^2/var.y^2)$
$print(estadistico.chi)$
$prob = pchisq(estadistico.chi, df = length(y) - 1)$
$print(prob)$

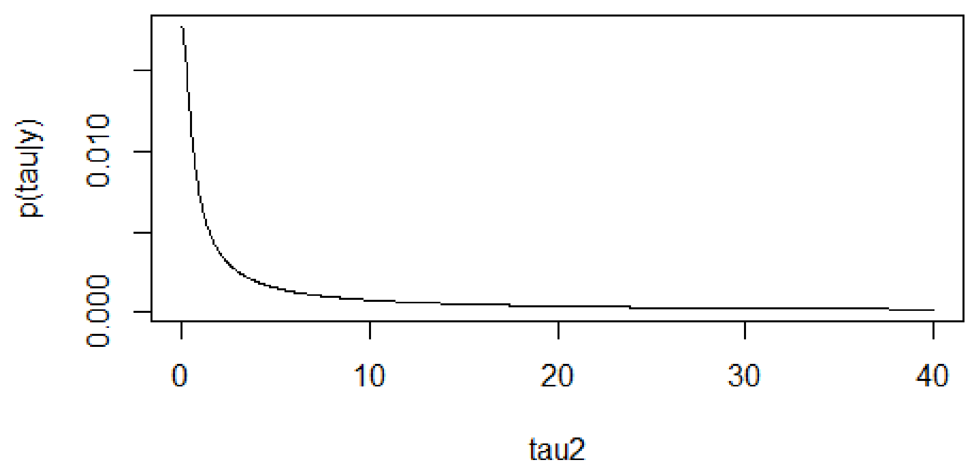
Se puede también verificar que variabilidad se esperaría si $\mu_1 = \mu_2$ mediante la distribución normal con media $\bar{y}_{..}$ y varianza la raíz cuadrada de la media de los s_i^2 en R
$test = rnorm(11*1000, ycombinado, sqrt(mean.var.y))quantile(test, c(,025, ,975))$
Lo que da que la variabilidad esperada sería $[-0{,}7168376, 1{,}1621442]$
La distribución de Pareto es una distribución de colas pesadas y podemos hacer una simulacion en grid para $[\tau_1, \tau 2]$ en $[1, 40]x[1, 40]$
Luego analizamos las distribuciones a posteriori de τ_1 y τ_2 con 1000 simulaciones y un grid de 2000 divisiones del intervalo [0,40] para τ_1 y τ_2. En R:
$n.sims = 1000$
$mu1.estimado < -function(tau1, y.controles, var.y.controles)\{$
$sum(y.controles/(var.y.controles^2 + tau1^2))/sum(1/(var.y.controles^2 + tau1^2))$
$\}$
$var.mu1.estimado < -function(tau1, y.controles,$
$var.y.controles)\{$
$1/sum(1/(tau1^2 + var.y.controles^2))$
$\}$
$mu2.estimado < -function(tau2, y.cohortes, var.y.cohortes)\{$

$sum(y.cohortes/(var.y.cohortes^2+tau2^2))/sum(1/(var.y.cohortes^2+tau2^2))$
$\}$
$var.mu2.estimado < -function(tau2, y.cohortes, var.y.cohortes)\{$
$1/sum(1/(tau2^2 + var.y.cohortes^2))$
$\}$
$n.grid < -2000$
$tau.grid < -seq(,01, 40, length = n.grid)$
$tau1.grid < -seq(,01, 40, length = n.grid)$
$tau2.grid < -seq(,01, 40, length = n.grid)$
$log.p.tau < -rep(NA, n.grid)$
$log.p.tau1 < -rep(NA, n.grid)$
$log.p.tau2 < -rep(NA, n.grid)$
$for(iin1 : n.grid)\{$
$mu1 < -mu1.estimado(tau1.grid[i], y.controles, var.y.controles)$
$V1 < -var.mu1.estimado(tau1.grid[i], y.controles, var.y.controles)$
$log.p.tau1[i] < -,5 * log(V1) -$
$,5 * sum(log(var.y.controles^2 + tau1.grid[i]^2)) -$
$,5 * sum((y.controles - mu1)^2/(var.y.controles^2 + tau1.grid[i]^2))$
$mu2 < -mu2.estimado(tau2.grid[i], y.cohortes, var.y.cohortes)$
$V2 < -var.mu2.estimado(tau2.grid[i], y.cohortes, var.y.cohortes)$
$log.p.tau2[i] < -,5 * log(V2) -$
$,5 * sum(log(var.y.cohortes^2 + tau2.grid[i]^2)) -$
$,5 * sum((y.cohortes - mu2)^2/(var.y.cohortes^2 + tau2.grid[i]^2))$
$\}$
$log.p.tau1 < -log.p.tau1 - max(log.p.tau1)$
$p.tau1 < -exp(log.p.tau1)$
$p.tau1 < -p.tau1/sum(p.tau1)$
$log.p.tau2 < -log.p.tau2 - max(log.p.tau2)$
$p.tau2 < -exp(log.p.tau2)$
$p.tau2 < -p.tau2/sum(p.tau2)$
$hist(p.tau1, probability = TRUE, breaks = 20, xlab = "tau1", ylab = "p(tau1|y)")$
$lines(density(p.tau1), col = "red")$
$hist(p.tau2, probability = TRUE, breaks = 20, xlab = "tau2", ylab = "p(tau2|y)")$
$lines(density(p.tau2), col = "red")$
$plot(tau1.grid, p.tau1, type = "l", xlab = "tau1", ylab = "p(tau|y)")$
$plot(tau2.grid, p.tau2, type = "l", xlab = "tau2", ylab = "p(tau|y)")$
Con lo que se obtiene


Histogram of p.tau1
p(tau1|y)
0
10
30
50
0.00
0.05
0.10
0.15
0.20
0.25
0.30
tau1

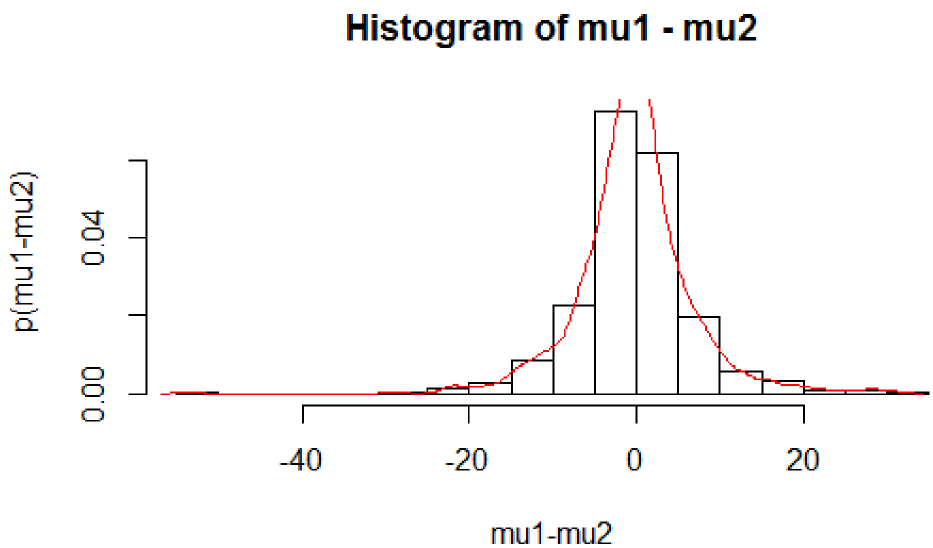

Histogram of p.tau2
p(tau2|y)
0
200
600
0.000
0.005
0.010
0.015
tau2





Para $p(\mu_1|y)$ y $p(\mu_2|y)$ podemos hacer una simulación en R extrayendo primero los valores de τ_1 y τ_2 repectivamente y entonces obtener la distrubución de $\mu_1|\tau_1, y$ y $\mu_2|\tau_2, y$. En R
$n.sims < -1000$

$tau1 < -sample(tau1.grid, n.sims, replace = T, prob = p.tau)$
$tau2 < -sample(tau2.grid, n.sims, replace = T, prob = p.tau)$
$mu1 < -rep(NA, n.sims)$
$mu2 < -rep(NA, n.sims)$
$theta < -array(NA, c(n.sims, J))$
$theta.mean < -numeric(n.sims)$
$theta.sd < -numeric(n.sims)$
$for(iin1 : n.sims)\{$
$mu1[i] < -rnorm(1, mu1.estimado(tau[i], y.controles, var.y.controles),$
$sqrt(var.mu1.estimado$
$(tau1[i], y.controles, var.y.controles)))$
$mu2[i] < -rnorm(1, mu2.estimado(tau[i], y.cohortes, var.y.cohortes),$
$sqrt(var.mu2.estimado$
$(tau2[i], y.cohortes, var.y.cohortes)))$
$\}$
$hist(mu1/mu2, breaks = 20, probability = TRUE, xlab = "mu1/mu2", ylab =$
$"p(mu1/mu2)")$
$hist(mu1{-}mu2, breaks = 20, probability = TRUE, xlab = "mu1{-}mu2", ylab =$
$"p(mu1 - mu2)")$
$lines(density(mu1 - mu2), col = "red")$
$mean(mu1 - mu2)$
$median(mu1 - mu2)$
$plot(mu1, mu2)$
$diferencia = (mu1 - mu2)/n.sims$
$print(sum(diferencia < 0))$
$quantile(diferencia, c(,025, ,975))$
lo que da un gráfico para $\mu_1 - \mu_2$



Con una media de $\mu_1 - \mu_2$ de -0.3602096 y un intervalo de confianza al 95 % de $[-14{,}99556, 13{,}38560]$.Como se ve la distribución es simétrica respecto a 0 y no evidencia para decir que $\mu_1 \neq \mu_2$