



Universidad Nacional de Colombia

FACULTAD DE CIENCIAS

PRIMER CASO DE ESTUDIO: VERIZON

Estadística Bayesiana

Docente:

Juan Camilo Sosa

Ander Steven Cristancho Sánchez (acristanchos@unal.edu.co)

Valeria del Pilar Lopez Bosiga (vlopezbo@unal.edu.co)

Septiembre de 2023

Introducción

Verizon es la principal compañía telefónica local (ILEC, incumbent local exchange carrier) para una gran área del este de Estados Unidos. Como tal, es responsable de brindar servicio de reparación a los clientes de otras empresas telefónicas de la competencia (CLEC, competing local exchange carrier) en esta región.

Verizon está sujeto a multas si los tiempos de reparación (el tiempo que lleva solucionar un problema) para los clientes de alguna CLEC son sustancialmente peores que los de los clientes de Verizon. El conjunto de datos Verizon.csv (Chihara and Hesterberg, 2019, Sec. 1.3) contiene una muestra de los tiempos de reparación de $n_1 = 1664$ clientes de Verizon (ILEC) y $n_2 = 23$ clientes de la competencia (CLEC). De acuerdo con los datos, los tiempos medios de reparación son $\bar{x}_1 = 8.41$ y $\bar{x}_2 = 16.51$ horas para ILEC y CLEC, respectivamente.

El objetivo de este caso es determinar si la diferencia entre los tiempos promedio de reparación es lo suficientemente grande para declararse como significativa, y por tanto, ser tomada en cuenta como evidencia para llevar a cabo una intervención y multar a Verizon.

Modelo

La distribución Exponencial es popular para modelar tiempos dado que este modelo permite producir distribuciones con diferentes tasas de decaimiento y variedades de sesgo (para más información acerca de este modelo probabilístico, ver por ejemplo Sosa et al. 2014, Sec. 10.4).

Así, considere modelos Exponenciales independientes de la forma

$$y_{k,i} \mid \lambda_k \stackrel{\text{iid}}{\sim} \text{Exp}(\lambda_k) \Leftrightarrow p(y_{k,i} \mid \lambda_k) = \frac{1}{\lambda_k} \exp\left(-\frac{y_{k,i}}{\lambda_k}\right), \quad y_{k,i} > 0, \quad \lambda_k > 0, \quad (1)$$

Para $i = 1, \dots, n_k$ y $k = 1, 2$ (1: ILEC, 2: CLEC), donde $y_{k,i}$ es el tiempo de reparación (en horas) del individuo i en el grupo k , n_k es el tamaño de la muestra del grupo k , y finalmente, $y_k = (y_{k,1}, \dots, y_{k,n_k})$ es el vector columna de observaciones correspondiente.

Análisis Bayesiano

1. Ajuste los modelos Gamma-Inversa-Exponencial con $a_k = 3$ y $b_k = 17$ en cada grupo. A partir de las distribuciones posteriores, obtenga la distribución posterior de η . Reporte la media, el coeficiente de variación y un intervalo de credibilidad al 95 % para η . Presente los resultados visual y tabularmente. Interprete los resultados obtenidos (máximo 100

palabras).

Solución:

Dada la distribución muestral 1 y la distribución previa de $\lambda_k \sim \text{GI}(\alpha, \beta)$, por el Teorema de Bayes, se tiene que,

$$\begin{aligned} p(\lambda_k | \mathbf{y}) &\propto p(\mathbf{y} | \lambda_k) \cdot p(\lambda_k) \\ &= \left[\lambda_k^{-n} \exp\left(-\frac{s_k}{\lambda_k}\right) \right] \cdot \left[\lambda_k^{-(\alpha+1)} \exp\left(-\frac{\beta}{\lambda_k}\right) \right] \\ &= \lambda_k^{-(\alpha+n+1)} \exp\left[-\frac{s_k + \beta}{\lambda_k}\right] \\ \lambda_k | \mathbf{y} &\sim \text{GI}(\alpha + n_k, \beta + s_k) \end{aligned}$$

Donde $s_k = \sum_{i=1}^{n_k} y_i$, para $k = 1, 2$. Teniendo en cuenta que ambos grupos tienen la misma distribución previa, las distribuciones posteriores correspondientes son:

$$\lambda_1 | \mathbf{y} \sim \text{GI}(\alpha + n_1, \beta + s_1) = \text{GI}(3 + 1664, 17 + 13996.92)$$

$$\lambda_2 | \mathbf{y} \sim \text{GI}(\alpha + n_2, \beta + s_2) = \text{GI}(3 + 23, 17 + 379.71)$$

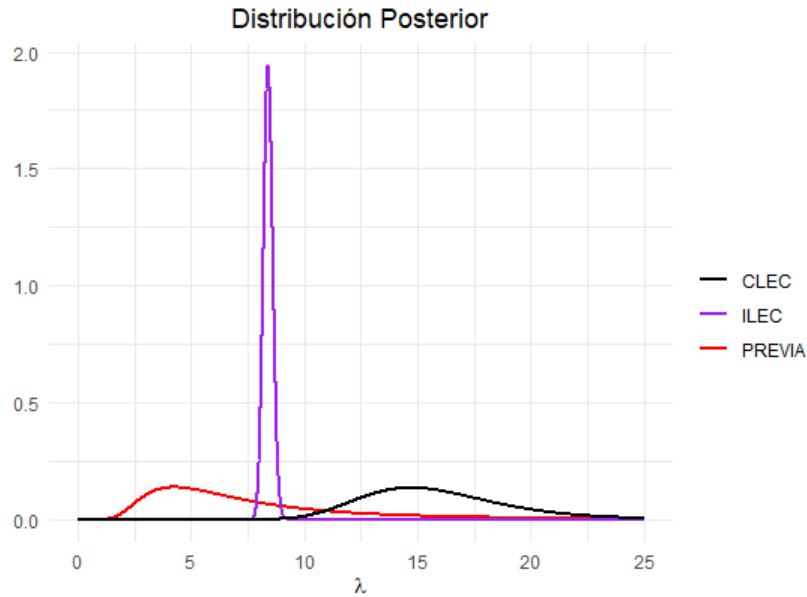


Figura 1: Distribución posterior por grupos vs. previa

Por otro lado, se tiene que $\eta = \lambda_1 - \lambda_2$. Se sabe que si $\eta = 0$ los tiempos promedio de reparación serían iguales, si $\eta < 0$ quiere decir que el tiempo medio de reparación de CLEC es mayor, mientras que si $\eta > 0$ el tiempo promedio de reparación ILEC es

mayor. Se usó método el de Monte Carlo con 10000 muestras para hallar la distribución posterior de η , en donde se obtuvo lo siguiente:

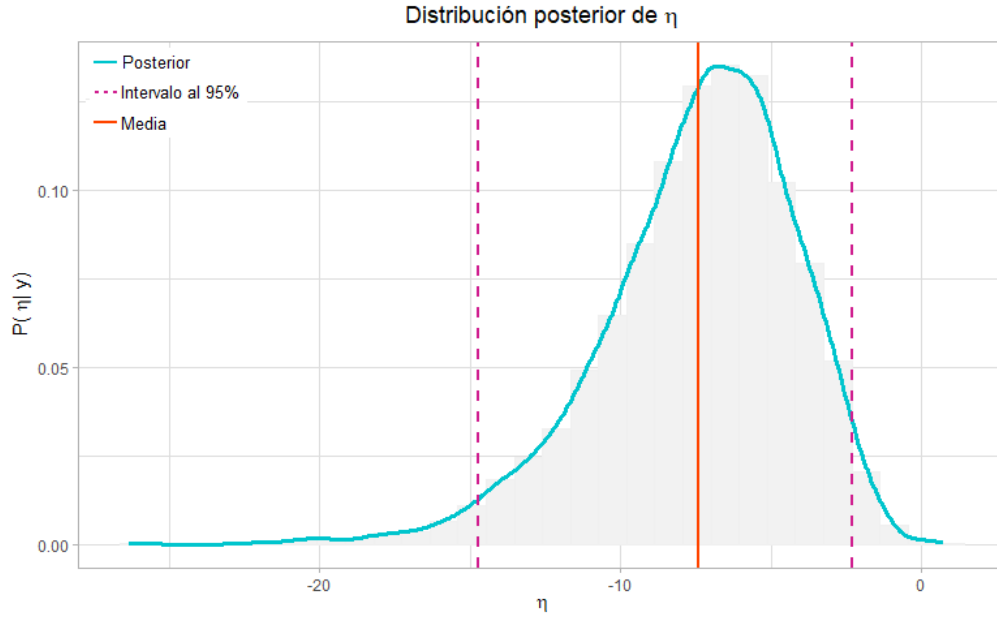


Figura 2: Distribución posterior de η

Media Posterior	CV Posterior	Q2.5 %	Q97.5 %
-7.429	0.432	-14.757	-2.333

Tabla 1: Estimaciones de η

A partir de esto, se puede afirmar con un 95% de probabilidad que el valor de η se encuentra contenido en el intervalo de $(-14.757, -2.333)$. Como este intervalo no contiene al cero y además toma solo valores negativos, se puede afirmar que es mayor el tiempo promedio de reparación en los usuarios de CLEC. Por otro lado, el coeficiente de variación es del 43.2% lo que indica alta variabilidad en la estimación. Por lo tanto, con la información recolectada, hay evidencia significativa para hacer una intervención y multar a Verizon.

2. Lleve a cabo un análisis de sensibilidad. Para ello, considere los siguientes estados de información externos al conjunto de datos:

- Distribución previa 1: $a_k = 3$ y $b_k = 17$, para $k = 1, 2$.
- Distribución previa 2: $a_k = 2$ y $b_k = 8.5$, para $k = 1, 2$.
- Distribución previa 3: $a_k = 3$ y $b_1 = 16.8$ y $b_2 = 33$, para $k = 1, 2$.

- Distribución previa 4: $a_k = 2$ y $b_1 = 8.4$ y $b_2 = 16.5$, para $k = 1, 2$.

En cada caso, calcule la media y el coeficiente de variación a priori, y repita el numeral anterior. Presente los resultados visual y tabularmente. Interprete los resultados obtenidos (máximo 100 palabras).

Solución:

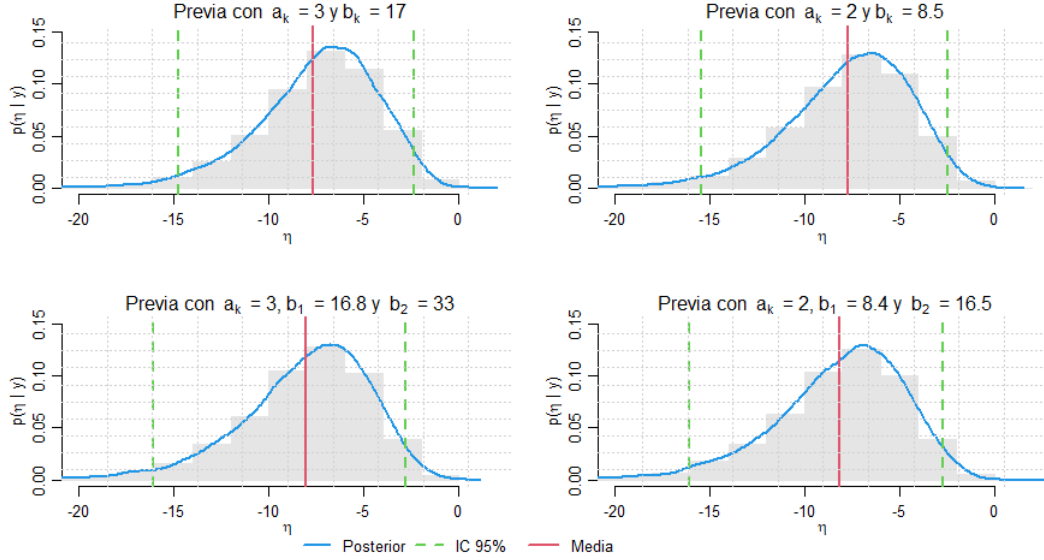


Figura 3: Análisis de sensibilidad

	Previa ILEC		Previa CLEC		Posterior		Int. Credebilidad	
Previa	Media	CV %	Media	CV %	Media	CV %	Q2.5 %	Q97.5 %
1	8.5	100	8.5	100	-7.429	43.184	-14.757	-2.333
2	8.5	indef.	8.5	indef.	-7.683	43.247	-15.393	-2.438
3	8.4	100	16.5	100	-8.072	42.172	-16.063	-2.795
4	8.4	indef.	16.5	indef.	-8.140	42.574	-16.085	-2.692

Tabla 2: Resultados del análisis de sensibilidad

A partir de los resultados presentados en la tabla 2, se aprecia la influencia de la información disponible sobre los parámetros en la distribución previa. Se observa que los coeficientes de variación de las previas son del 100 % o están indefinidos, lo que indica que son poco informativas. Los coeficientes de variación indefinidos se deben a que en esas previas se usó un hiperparámetro $\alpha_k = 2$, lo que lleva a que la varianza no esté definida para dichos casos. Por otro lado, al hacer el contraste con la distribución posterior, se puede observar que esta es sensible ante la información proporcionada por la distribución previa. Los valores de la media van cambiando a medida que los hiperparámetros varían,

lo que puede indicar que el conocimiento externo afecta a la distribución posterior y sus estimaciones. Sin embargo, el cambio en la inferencia posterior no es sustancial, pues las conclusiones a las que se habrían llegado son las mismas que con cualquiera de las distribuciones previas.

3. En cada población, evalúe la bondad de ajuste del modelo propuesto utilizando la distribución previa 1, utilizando como estadísticos de prueba la media y la desviación estándar. Presente los resultados visual y tabularmente. Interprete los resultados obtenidos (máximo 100 palabras).

Nota: Calcule los valores predictivos posteriores y, en cada grupo, realice la visualización de las distribuciones predictivas de los estadísticos de prueba de manera conjunta (dispersograma con histogramas marginales).

Solución:

Teniendo en cuenta a la media y la desviación estándar como los estadísticos de prueba, se calcularon los valores predictivos posteriores para ILEC y CLEC. Se llevaron a cabo métodos de Monte Carlo para evaluar la bondad de ajuste del modelo, en donde se obtuvo lo siguiente:

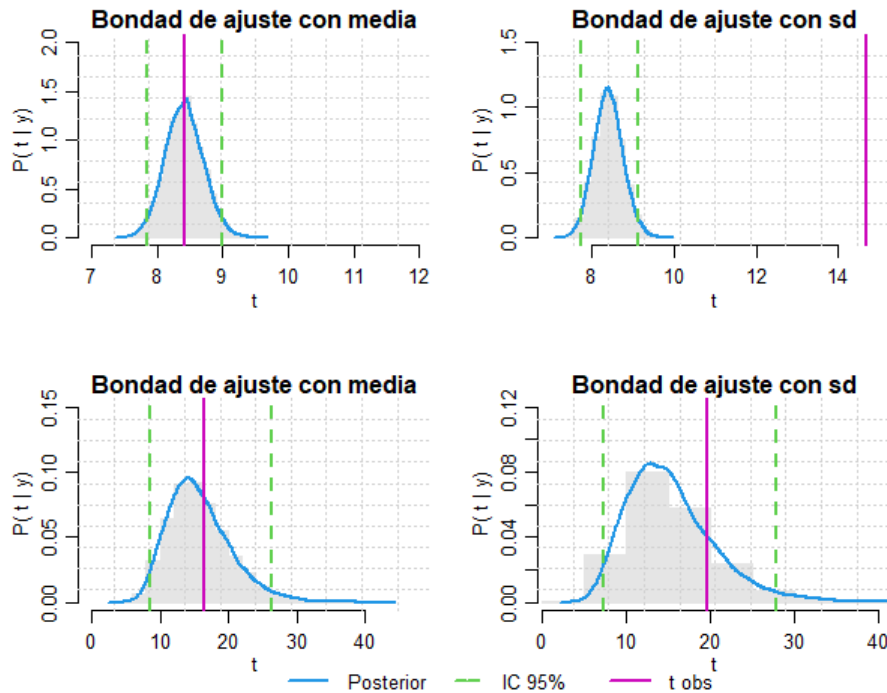


Figura 4: Bondad de ajuste

Estadístico	PPP	t observado	Q 2.5 %	Q 97.5 %
Media ILEC	0.502	8.412	7.855	8.997
Media CLEC	0.385	16.509	8.586	26.434
Desviación estándar ILEC	0	14.690	7.729	9.127
Desviación estándar CLEC	0.185	19.504	7.330	27.792

Tabla 3: Bondad de ajuste

A partir de los resultados obtenidos en la tabla 3, se puede afirmar que el modelo captura adecuadamente la media para ILEC y CLEC, dado a que el estadístico correspondiente para cada población se encuentra dentro del intervalo de credibilidad del 95 %. Sin embargo, la desviación estándar no es capturada por el modelo en la población de ILEC, ya que su ppp es igual a cero, mientras que logra ser capturada para la población de CLEC en donde su estadístico correspondiente se encuentra dentro de los límites del intervalo de credibilidad.

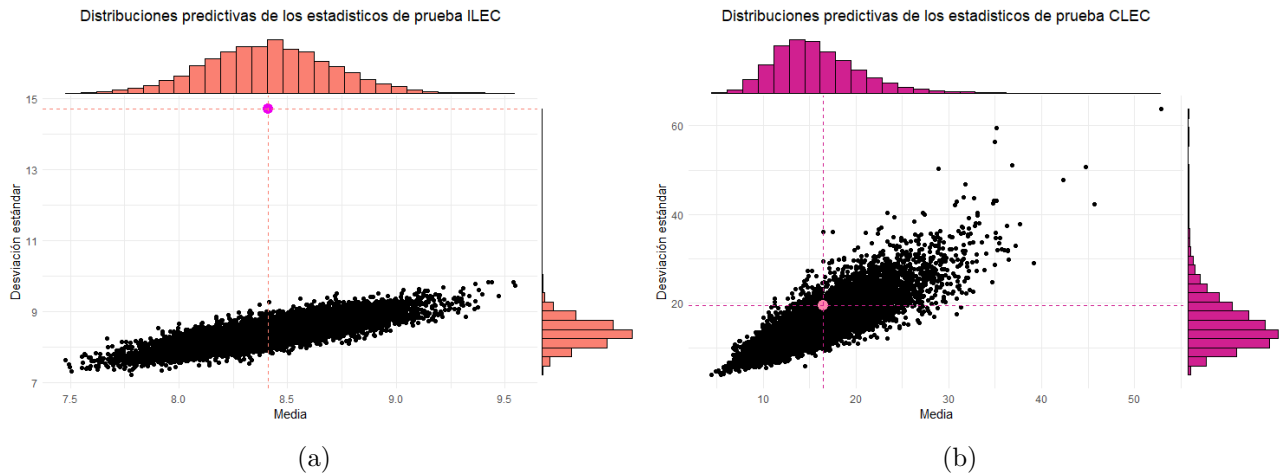


Figura 5: Distribuciones predictivas de los estadísticos de prueba de manera conjunta

Análisis frecuentista

Repita el numeral 1. de la PARTE 1 usando la Normalidad asintótica del MLE, *bootstrap* paramétrico y *bootstrap* no paramétrico. Presente los resultados visual y tabularmente. Interprete los resultados obtenidos (máximo 100 palabras).

Nota: Asintóticamente se tiene que $\hat{\lambda}_{MLE} \approx N(\lambda, \hat{I}^{-1})$, donde $\hat{\lambda}_{MLE}$ es el MLE de λ y \hat{I} es la información observada de Fisher.

Nota: Cuando utilice *bootstrap*, use una cantidad adecuada de remuestras y el método de los percentiles para calcular los intervalos de confianza.

Solución:

- **Distribución asintótica del MLE:** Para estimar $\eta = \lambda_1 - \lambda_2$ utilizando la normalidad asintótica del MLE, primero se tiene que ver la distribución del MLE de λ_k .

Recordando el modelo exponencial independiente de ambas muestras, tenemos que la función de densidad del modelo es:

$$p(y_k) = \frac{1}{\lambda_k} \exp\left(-\frac{y_k}{\lambda_k}\right) I_{(0,\infty)}(y_k) \quad (2)$$

Y esto se puede ver como:

$$a(\lambda_k)b(y_k)\exp[c(\lambda_k)d(y_k)]$$

Donde

$$a(\lambda_k) = \frac{1}{\lambda_k}, b(y_k) = I_{(0,\infty)}(y_k), c(\lambda_k) = -\frac{1}{\lambda_k}, d(y_k) = y_k \quad (3)$$

Luego, según la definición 2.2.36 de Mayorga[1], se tiene que la distribución exponencial pertenece a la familia exponencial de densidades uniparamétrica.

Sabiendo esto, se halla un estadístico suficiente para λ_k y su MLE.

$$\sum_{i=1}^{n_k} d(y_{k,i}) = \sum_{i=1}^{n_k} y_{k,i} = s_k \quad (4)$$

Luego s_k es un estadístico suficiente para λ_k .

Además, se tiene que el estimador máximo verosímil de λ se puede hallar resolviendo la siguiente ecuación:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{n_k} d(Y_{k,i}) &= E[d(Y_k)]|_{\lambda_k = \hat{\lambda}_{kMLE}} \\ \frac{1}{n} \sum_{i=1}^{n_k} Y_{k,i} &= E[Y_k]|_{\lambda_k = \hat{\lambda}_{kMLE}} \\ \bar{Y}_k &= \hat{\lambda}_{kMLE} \end{aligned} \quad (5)$$

Luego, el estimador MLE de λ_k es $\hat{\lambda}_{kMLE} = \bar{Y}_k$

Ya que $\hat{\lambda}_{kMLE} \xrightarrow{n \rightarrow \infty} N(\lambda, \hat{I}^{-1})$, se calcula la información observada de Fisher \hat{I} para tener la distribución asintótica de $\hat{\lambda}_{kMLE}$:

$$\hat{I}(\hat{\lambda}_{kMLE}) = -\frac{\partial^2}{\partial \lambda^2} \ln L(\lambda|y) \Bigg|_{\lambda = \hat{\lambda}_{kMLE}} \quad (6)$$

Primero se desarrolla la logverosimilitud y luego se deriva:

$$\begin{aligned}
L(\lambda_k | \mathbf{y}_k) &= \prod_{i=1}^{n_k} p(\lambda | y_{k,i}) \\
&= \prod_{i=1}^{n_k} \frac{1}{\lambda_k} \exp\left(-\frac{y_{k,i}}{\lambda_k}\right) I_{(0,\infty)}(y_{k,i}) \\
&= \frac{1}{\lambda_k^{n_k}} \exp\left(-\frac{\sum_{i=1}^{n_k} y_{k,i}}{\lambda_k}\right) \prod_{i=1}^{n_k} I_{(0,\infty)}(y_{k,i})
\end{aligned} \tag{7}$$

$$\begin{aligned}
\ln L(\lambda_k | \mathbf{y}_k) &= \ln \left[\frac{1}{\lambda_k^{n_k}} \exp\left(-\frac{\sum_{i=1}^{n_k} y_{k,i}}{\lambda_k}\right) \prod_{i=1}^{n_k} I_{(0,\infty)}(y_{k,i}) \right] \\
&= -n_k \ln(\lambda_k) - \frac{s_k}{\lambda_k} + \ln \left(\prod_{i=1}^{n_k} I_{(0,\infty)}(y_{k,i}) \right)
\end{aligned} \tag{8}$$

$$\begin{aligned}
\frac{\partial}{\partial \lambda} \ln L(\lambda | y) &= \frac{\partial}{\partial \lambda} \left[-n_k \ln(\lambda_k) - \frac{s_k}{\lambda_k} + \ln \left(\prod_{i=1}^{n_k} I_{(0,\infty)}(y_{k,i}) \right) \right] \\
&= -\frac{n_k}{\lambda_k} + \frac{s_k}{\lambda_k^2}
\end{aligned} \tag{9}$$

$$\begin{aligned}
-\frac{\partial^2}{\partial \lambda^2} \ln L(\lambda | y) \Big|_{\lambda=\hat{\lambda}_{kMLE}} &= -\frac{\partial}{\partial \lambda} \left[-\frac{n_k}{\lambda_k} + \frac{s_k}{\lambda_k^2} \right] \Big|_{\lambda=\hat{\lambda}_{kMLE}} \\
&= -\frac{n_k}{\lambda_k^2} + \frac{2s_k}{\lambda_k^3} \Big|_{\lambda=\hat{\lambda}_{kMLE}} \\
&= -\frac{n_k}{(s_k/n_k)^2} + \frac{2s_k}{(s_k/n_k)^3} \\
&= -\frac{n_k^3}{s_k^2} + \frac{2n_k^3}{s_k^2} = \frac{n_k^3}{s_k^2} = \frac{n_k}{\bar{y}_k^2}
\end{aligned} \tag{10}$$

Con esto, se tiene que $\hat{\lambda}_{kMLE} \xrightarrow{n \rightarrow \infty} N(\lambda, \frac{\bar{y}_k^2}{n_k})$.

Ahora, se quiere hacer inferencia sobre $\eta = \lambda_1 - \lambda_2$, por lo que, soportado por el teorema 2.1.12 de Mayorga[1], que hace referencia a la invarianza del MLE, se tiene que $\hat{\eta}_{MLE} = \hat{\lambda}_{1MLE} - \hat{\lambda}_{2MLE}$. Y además, es sabido que la resta de 2 variables aleatorias con distribución normal se distribuye normal con media la diferencia de las medias y varianza la suma de las varianzas, por lo que $\hat{\eta}_{MLE} \xrightarrow{n \rightarrow \infty} N\left(\lambda_1 - \lambda_2, \frac{\bar{y}_1^2}{n_1} + \frac{\bar{y}_2^2}{n_2}\right)$,

Teniendo esto se puede hacer inferencia sobre $\hat{\eta}_{MLE}$, por lo que los estadísticos que son de interés son:

- **Media:** En este caso, la estimación puntual de η es dada por el estimador máximo

verosímil, por lo que:

$$\hat{\eta}_{MLE} = \bar{y}_1 - \bar{y}_2 \quad (11)$$

- **Coeficiente de variación:** En este caso, se puede calcular el coeficiente de variación con la estimación de la varianza y de la media de $\hat{\eta}_{MLE}$, por lo que se tiene que:

$$CV(\hat{\eta}_{MLE}) = \left| \frac{\sqrt{\frac{\bar{y}_1^2}{n_1} + \frac{\bar{y}_2^2}{n_2}}}{\hat{\lambda}_{1MLE} - \hat{\lambda}_{2MLE}} \right| = \frac{\sqrt{\frac{\bar{y}_1^2}{n_1} + \frac{\bar{y}_2^2}{n_2}}}{|\bar{y}_1 - \bar{y}_2|} \quad (12)$$

- **Intervalo de confianza del 95 %:** Los límites del intervalo vienen dados por:

$$IC_{95\%}(\eta) = \hat{\eta}_{MLE} \pm z_{97.5} \sqrt{\frac{\bar{y}_1^2}{n_1} + \frac{\bar{y}_2^2}{n_2}} \quad (13)$$

Al no conocer λ_1 y λ_2 , se imposibilita graficar la densidad de $\hat{\eta}_{MLE}$, sin embargo, se puede graficar una aproximación de cómo se vería la distribución de $\hat{\eta}_{MLE}$, tomando como media a $\hat{\lambda}_{1MLE} - \hat{\lambda}_{2MLE}$, el gráfico es el siguiente:

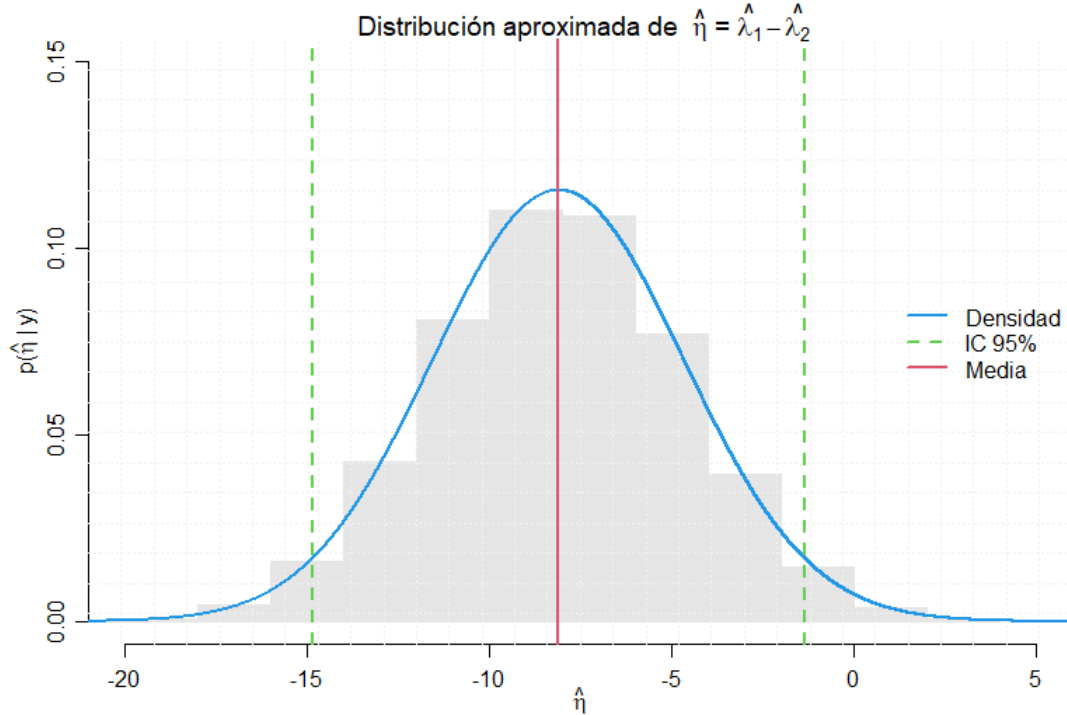


Figura 6: Distribución asimptótica de $\hat{\eta}_{MLE}$

- **Bootstrap paramétrico:** Para estimar η a través de *bootstrap* paramétrico, se generaron $B = 10000$ muestras para cada grupo, cada muestra de tamaño n_k y con parámetro

de la distribución \bar{y}_k , que como vimos en 5, es el estimador máximo verosímil de λ_k , para $k = 1, 2$. Luego se calculó \bar{y} para cada muestra de ambos grupos, y con esta información conseguimos el vector $\hat{\boldsymbol{\eta}} = (\hat{\eta}_1, \hat{\eta}_2, \dots, \hat{\eta}_B)$, donde $\hat{\eta}_j = \bar{y}_{j,1} - \bar{y}_{j,2}$ con $j = 1, 2, \dots, B$. Este vector sería una aproximación a la distribución de $\hat{\eta}_{MLE}$, la cual puede verse en el siguiente gráfico:

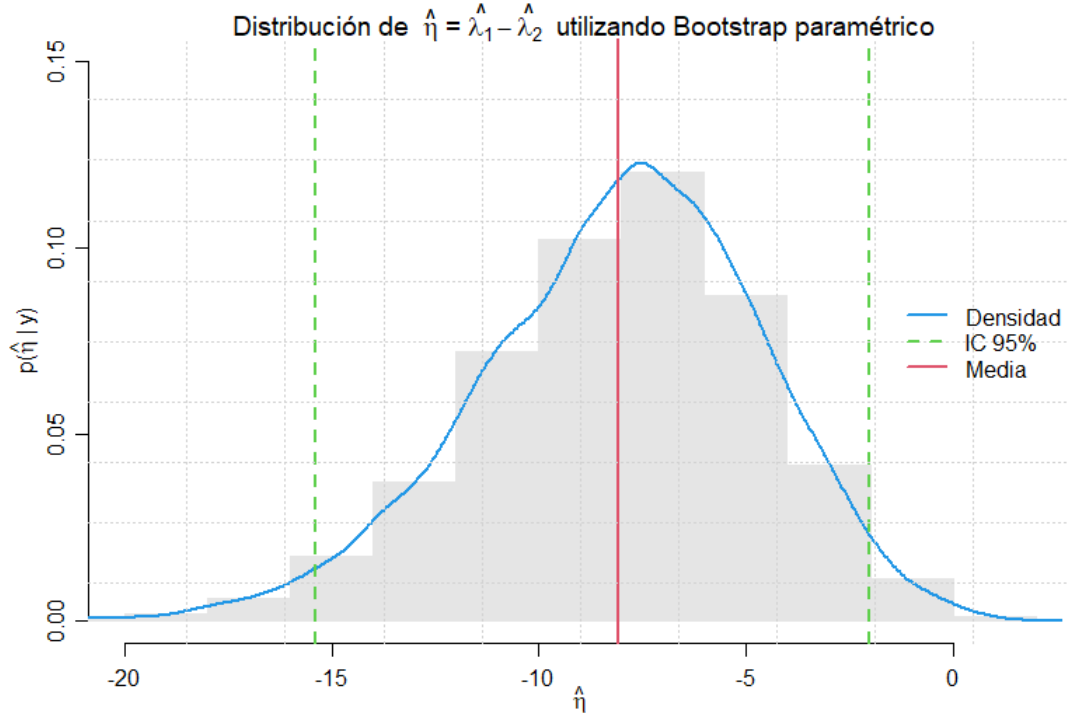


Figura 7: *Bootstrap* paramétrico

- ***Bootstrap* no paramétrico:** Para estimar η a través de *Bootstrap* no paramétrico, se hicieron $B = 10000$ remuestras para cada grupo, cada muestra de tamaño n_k y con reemplazo, a partir de la muestra original de cada grupo. Luego se calculó \bar{y} para cada muestra de ambos grupos, y con esta información conseguimos el vector $\hat{\boldsymbol{\eta}} = (\hat{\eta}_1, \hat{\eta}_2, \dots, \hat{\eta}_B)$, donde $\hat{\eta}_j = \bar{y}_{j,1} - \bar{y}_{j,2}$ con $j = 1, 2, \dots, B$. Este vector sería una aproximación a la distribución de $\hat{\eta}_{MLE}$, y gráficamente es:

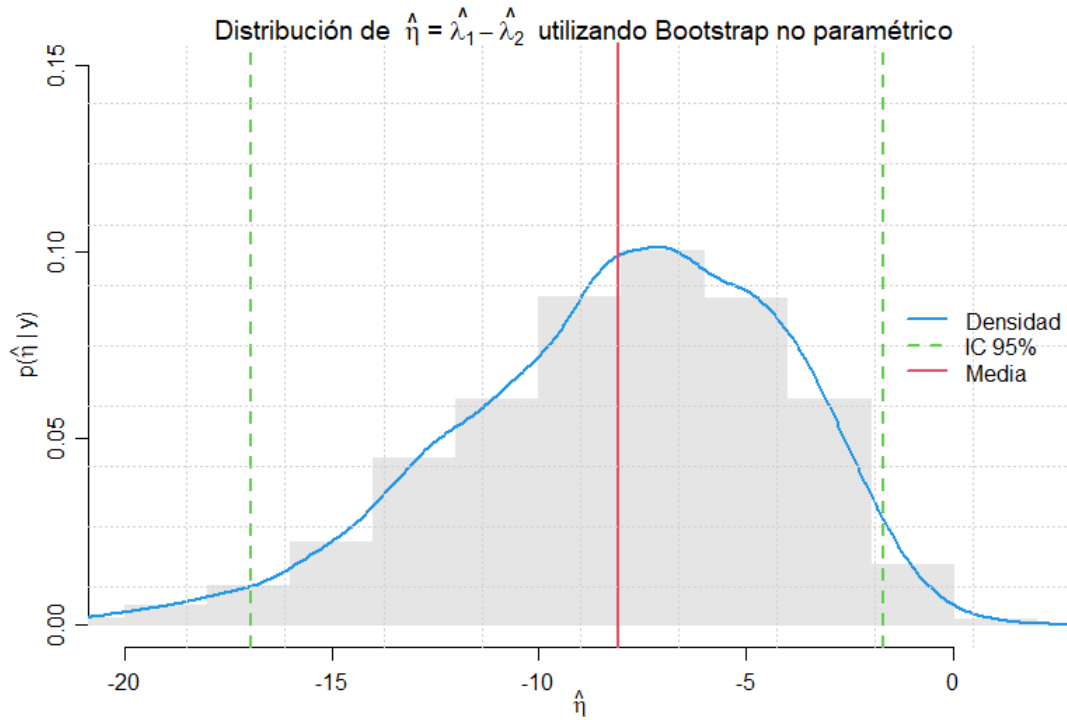


Figura 8: *Bootstrap* no paramétrico

Las estimaciones de la media, el coeficiente de variación y los límites del intervalo de confianza al 95 %, para cada método son:

Método	Media	CV(%)	Q2.5 %	Q97.5 %
Normalidad asintótica	-8.098	42.6	-14.857	-1.338
<i>Bootstrap</i> paramétrico	-8.075	42.3	-15.381	-2.016
<i>Bootstrap</i> no paramétrico	-8.096	42.3	-16.970	-1.691

Tabla 4: Resultados de inferencia frecuentista sobre η

Se puede apreciar que las estimaciones de la media y los coeficientes de variación son muy parecidos entre sí, aunque la diferencia más notoria se encuentra en los límites de los intervalos de cada método, siendo el intervalo del *bootstrap* no paramétrico el que tiene una mayor amplitud, lo cual tiene sentido al ser el resultado de un método no paramétrico. Pero, además de esto, no hay diferencias significativas entre los resultados que nos lleven a pensar que los resultados de un método priman sobre otro. Por lo que siguiendo este análisis, hay evidencia significativa para multar a Verizon.

Respecto al análisis bayesiano, se llegan a las mismas conclusiones, sin embargo, las estimaciones difieren por muy poco, ya que las estimaciones de η en este caso son más pequeñas, o lo mismo, la diferencia de tiempos de reparación estimada es un poco mayor

para el análisis frecuentista. Sin embargo, en el caso de coeficientes de variación no se logra ver una diferencia significativa.

Simulación

Simule 100,000 muestras aleatorias de poblaciones Exponenciales bajo los siguientes escenarios:

- Escenario 1: $n_1 = 10, n_2 = 10, \lambda_1 = \bar{y}_1$, y $\lambda_2 = \bar{y}_2$.
- Escenario 2: $n_1 = 20, n_2 = 20, \lambda_1 = \bar{y}_1$, y $\lambda_2 = \bar{y}_2$.
- Escenario 3: $n_1 = 50, n_2 = 50, \lambda_1 = \bar{y}_1$, y $\lambda_2 = \bar{y}_2$.
- Escenario 4: $n_1 = 100, n_2 = 100, \lambda_1 = \bar{y}_1$, y $\lambda_2 = \bar{y}_2$.

Donde $\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} y_{k,i}$ es la media muestral observada del grupo k . Observe que el valor verdadero de η en cada caso es $\eta = \lambda_1 - \lambda_2 = \bar{y}_1 - \bar{y}_2$.

Usando cada muestra, ajuste el modelo de manera tanto Bayesiana (usando la distribución previa 1) como frecuentista (usando la Normalidad asintótica, *bootstrap* paramétrico, *bootstrap* no paramétrico), y en cada caso calcule la proporción de veces que el intervalo de credibilidad/confianza al 95 % contiene el valor verdadero de η . Reporte los resultados tabularmente. Interprete los resultados obtenidos (máximo 100 palabras).

Solución: Se simularon $K = 100000$ muestras aleatorias para las poblaciones $k = 1, 2$, en cuatro escenarios, cada escenario con un tamaño de muestra diferente, $n_k = 10, 20, 50, 100$, provenientes de una distribución exponencial con parámetro $\lambda_k = \bar{y}_k$. El verdadero valor del parámetro η es $\eta = -8.1$, y con las K muestras se realizaron intervalos de confianza utilizando cuatro métodos distintos. Para el método bayesiano se utilizaron simulaciones de Montecarlo con $B = 10000$ y para los métodos frecuentistas, *bootstrap* paramétrico y *bootstrap* no paramétrico, se hicieron $N = 1000$ remuestras. Luego se calcularon los intervalos de confianza en cada muestra, con el fin de ver si estos intervalos contenían o no al verdadero valor del parámetro η . Y por último, se vio la proporción de intervalos que sí contenían al verdadero valor del parámetro, consiguiendo así la cobertura expresada en (%) para cada método. Los resultados para los cuatro métodos son:

n_k	Bayesiano	N. asintótica	Bp. paramétrico	Bp. no paramétrico
10	93.278	94.648	93.862	89.123
20	93.921	94.869	94.300	91.701
50	94.513	94.930	94.614	93.506
100	94.768	95.032	94.720	94.192

Los resultados muestran que el intervalo de confianza a través de *bootstrap* no paramétrico es el que tiene peor desempeño, pues es el que más se aleja de nuestro nivel de confianza del 95 %, y que depende mucho de un tamaño de muestra grande, pues la cobertura se acerca considerablemente al nivel de confianza al aumentar el tamaño de muestra. El enfoque bayesiano y el *bootstrap* paramétrico se desempeñaron muy bien en general, en este caso fue mejor el *bootstrap* paramétrico, excepto con $n_k = 100$, pero puede deberse a la semilla utilizada, además de que computacionalmente es mucho más costoso el bootstrap. El que mejor desempeño tuvo fue el intervalo de confianza utilizando la normalidad asintótica del MLE, pues ya que con tamaños de muestra pequeños la cobertura estaba muy cerca al nivel de confianza.

Referencias

- [1] Mayorga Álvarez, J.(2004). *Inferencia estadística*. Universidad Nacional de Colombia