Construcción de Modelo Bayesiano de datos de Remanufactura

Dr. Manuel Arnoldo Rodríguez Medina[1], [1]MII Luz Isaura Rodríguez Aguilar[1], Dr. Manuel Alonso Rodríguez Morachis[1], Dr. Manuel Iván Rodríguez Borbón[2]

Resumen - Los productos regresados por los clientes para remanufactura presentan una gran incertidumbre con respecto a la calidad, lo cual se traduce en una significante variabilidad, complicando de gran manera el control de los inventarios. Es importante aquí diferenciar los procesos de reparación de los procesos de remanufactura, donde los primeros se dedican a reparar el daño o defecto del producto, mientras que la remanufactura incluye un análisis de todos los componentes del producto. La incertidumbre puede ser modelada mediante el análisis del comportamiento de cada una de las variables incluidas y construyendo una distribución a posteriori usando análisis Bayesiano.

Palabras Clave: Remanufactura, Análisis Bayesiano, Distribución a priori, Distribución a posteriori, Modelo Bayesiano

Introducción

Una de las opciones con la que es posible recuperar o crear valor a los productos que ya cumplieron con un ciclo de vida útil es la remanufactura. La remanufactura es el proceso de regresar un producto usado a un estado funcional de nuevo. En el proceso de remanufactura se recupera una proporción sustancial del recurso incorporado al producto durante su primera manufacturación. Lund (1984) define la remanufactura como la restauración de los productos usados a condiciones perfectas, proporcionándoles las características de rendimiento y durabilidad como las del producto original. La Figura 1 esquematiza de una manera básica el proceso de remanufactura

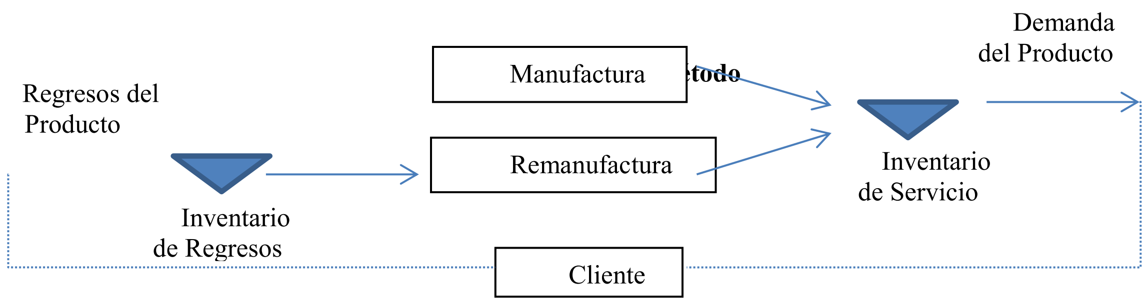

Figura 1 Esquema básico de un proceso de remanufactura

En cada uno de los procesos incluidos en la remanufactura existe gran incertidumbre, lo que implica la importancia de incluir formas de análisis del comportamiento que reduzcan la probabilidad de hacer inferencias incorrectas. El análisis Bayesiano es una metodología adecuada para hacer inferencias cuando existe incertidumbre.

Inferencia Bayesiana

El Teorema de Bayes se puede expresar de la manera siguiente:

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)} \tag{1}$$

donde $P(\theta|y)$ es la distribución a posteriori, $P(y|\theta)$ es la función de verosimilitud y $P(\theta)$ es la función a priori de los parámetros, en este caso, $(\mu,\ \sigma^2)$ de la distribución normal. El producto de las funciones $P(\theta|y) * P(\theta)$ constituye la distribución conjunta, mientras que el denominador $P(y)$ es la distribución marginal.

[1] Dr. Manuel Arnoldo Rodríguez Medina es Profesor Investigador de la DEPI del Instituto Tecnológico de Cd. Juárez, manuel_rodriguez_itcj@yahoo.com
[1]MII Luz Isaura Rodríguez Aguilar es estudiante de Doctorado en Ciencias en Ingeniería Industrial de la DEPI del ITCJ, luz_rodriguez10@yahoo.com.mx
[1]Dr. Manuel Alonso Rodríguez Morachis es Profesor Investigador de la DEPI del Instituto Tecnológico de Cd. Juárez, mmorachis@itcj.edu.mx
[2]Dr. Manuel Iván Rodríguez Borbón es Profesor Investigador del Instituto de Ingeniería y tecnología de la Universidad Autónoma de Cd. Juárez, ivan.rodriguez@uacj.mx

Los componentes en el modelo de Bayes serán fundamentados, es decir, las distribuciones a priori para los parámetros, la función de verosimilitud, la distribución marginal y la distribución a posteriori, con la cual se [i]desarrollarán las estimaciones antes mencionadas. Los cálculos se harán mediante el lenguaje R para cálculos estadísticos. Enseguida se incluyen los conceptos fundamentales para la obtención de la distribución a posteriori.[ii]

La distribución normal está dada por

$$f(y|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \quad (2)$$

y de acuerdo a Box&Tiao (1973), considerando que $\bar{y}$ se dice ser un estadístico suficiente para μ, y asumiendo de acuerdo a (Box&Tiao (1973), Albert (2009)), además de Gelman (2004), la a priori estándar no informativa está dada por

$$g(\mu,\sigma^2|y) \propto \frac{1}{\sigma^2} \quad (3)$$

y la función de verosimilitud se puede escribir

$$l(\mu,\sigma|y) \propto \sigma^{-n} \exp\left\{-\frac{1}{\sigma^2}[(n-1)s^2] + n(\bar{y}-\mu)^2\right\} \quad (4)$$

donde el producto de (3) y (4) forman la distribución conjunta. Esto es,

$$l(\mu,\sigma|y)\cdot\ g(\mu,\sigma^2|y) \propto \sigma^{-(n+2)} \exp\left\{-\frac{1}{\sigma^2}[(n-1)s^2] + n(\bar{y}-\mu)^2\right\} \quad (5)$$

La Ecuación (5) dará forma a la distribución a posteriori en el Teorema de Bayes, es decir, $P(\theta|y)$.

Las variables del proceso de remanufactura que se incluyen en el modelo son:

- Unidades procesadas esperadas por unidad de tiempo
- Unidades aceptadas esperadas por unidad de tiempo
- Unidades de nuevos productos esperados por unidad de tiempo
- Tasa de reutilización esperada por unidad de tiempo

y la forma de introducirlas en el lenguaje R construyendo un marco de datos para su análisis es:

```
> data=data.frame(PROCESSED,ACCEPTED,RATIO,NP)
> data
```

La Tabla 1 muestra la construcción de la tabla de datos como resultado del data.frame.

Tabla 1 Número de piezas probadas, aceptadas, tasa de reutilización y nuevos productos

Sem ana	PROCES SED	ACCEP TED	RATIO	NP
1	48541	21951	0.45222	3194
2	51103	34233	0.66988	2975
3	38590	23885	0.61894	3934
4	27011	18545	0.68657	3786
5	39488	28757	0.72825	5678
6	35660	24053	0.67451	4692
7	26484	15803	0.59670	4998
8	31690	22668	0.71530	3672
9	32699	22843	0.69858	3825
10	35292	25085	0.71078	4237
11	32076	26348	0.82142	4132
12	33831	27050	0.79956	4672

El siguiente programa en R está diseñado para la construcción de histogramas con curva Normal. Como se observa en los histogramas construidos (Figura 1) es evidente que las variables se comportan normalmente. Lo anterior es comprobado posteriormente por medio de una prueba Anderson-Darling.

```
x=PROCESSED
> h<-hist(x,breaks=12,xlab="PROCESSED",main="Histograma con curva normal")
>  xfit<-seq(min(x),max(x),length=20)
>  yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
>  yfit<-yfit*diff(h$mids[1:2])*length(x)
>  lines(xfit,yfit,col="blue",lwd=2)
```

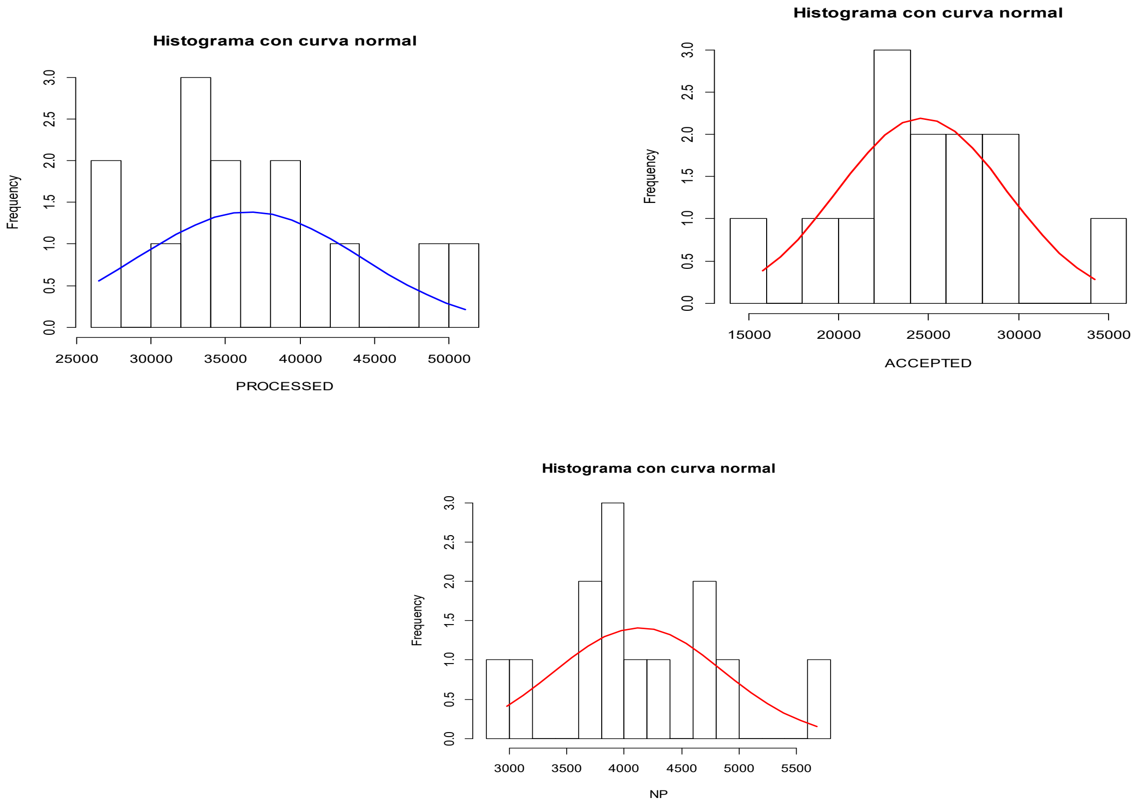

Figura 2 Histogramas de productos procesados, aceptados y nuevos productos.

Los valores p en la Tabla 2 demuestran de una manera analítica el comportamiento Normal de las variables incluidas en el análisis (Procesadas, aceptadas y nuevos productos). La comprobación de normalidad permite establecer la función de verosimilitud incluida en el Teorema de Bayes; es decir, asumir la función de densidad Normal como función de verosimilitud. Esta función escrita de manera conjunta con una previa no informativa, formaran la distribución conjunta en el Teorema de Bayes.

Tablas 2 Pruebas de Anderson Darling de Normalidad para los procesados, aceptados y Nuevos Productos

Probando Normalidad

> ad.test(PROCESSED) Anderson-Darling normality test data: PROCESSED A = 0.33625, p-value = 0.4483	> ad.test(ACCEPTED) Anderson-Darling normality test data: ACCEPTED A = 0.19478, p-value = 0.8645	>ad.test(NP) Anderson-Darling normality test data: NP A = 0.26356, p-value = 0.6371

La Figura 3 muestra el comportamiento de los datos en la distribución a posteriori, mediante gráficos de contorno con los valores de varianza en función de los valores de la media. Las Figura 4 muestra el comportamiento de la media aritmética y las varianza de las piezas procesadas. Estos valores fueron generados de distribuciones a posteriori simuladas mediante la generación de 1000 valores.

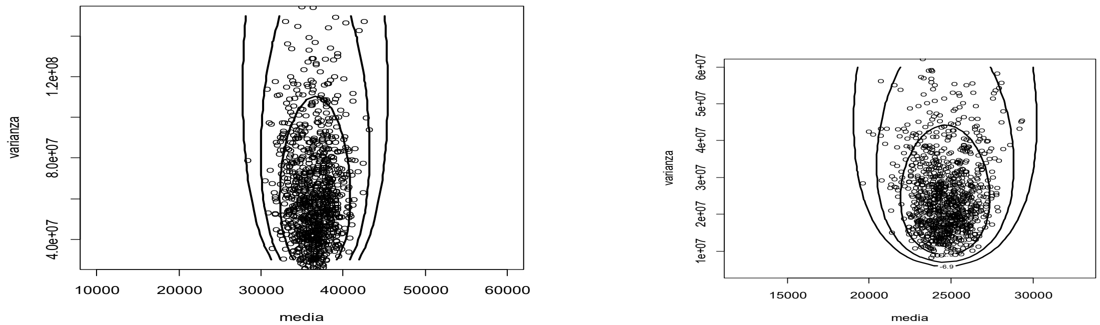

Figura 3 Gráficas de Contorno describiendo el comportamiento de las piezas procesadas y las aceptadas

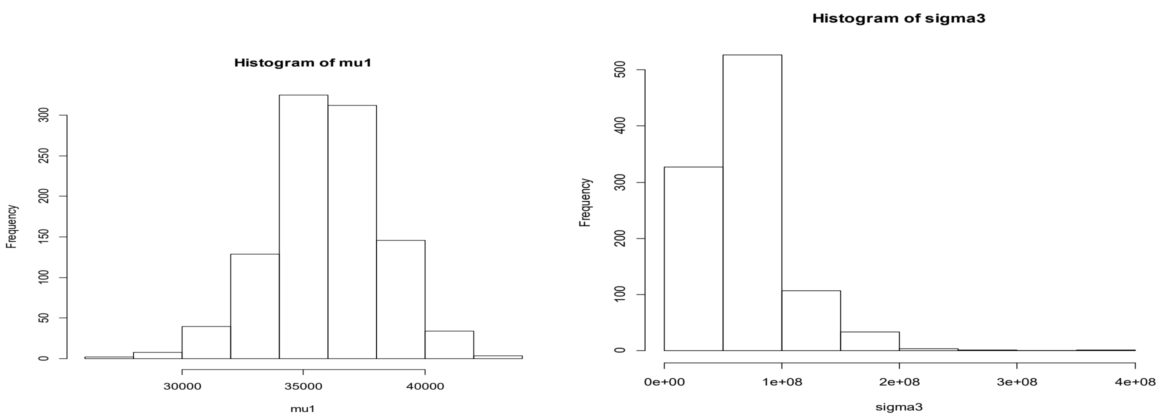

Figura 4 Comportamiento de la media aritmética y la varianza de las piezas procesadas (mu1 y Sigma3)

El Modelo de Regresión Lineal

Albert (2009) menciona que, generalmente, al construir un modelo de regresión, estamos interesados en describir la variación de la variable de respuesta y en términos de k variables predictoras. Se puede describir el valor promedio de y_i, la respuesta para el i-ésimo individuo, como

$$E(y_i|\beta, X) = \beta_1 x_{i1} + \cdots + \beta_k x_{ik}, i = 1, \ldots, n \qquad (6)$$

donde las $x_{i1}, \ldots x_{ik}$ son los valores predictores para el i-ésimo individuo y $\beta_1, \ldots, \beta_k$ son parámetros de regresión desconocidos. El modelo también puede ser representado mediante un vector renglón de variables regresoras y un vector columna de coeficientes de regresión como

$$E(y_i|\beta, X) = x_i\beta \qquad (7)$$

Donde las $\{y_i\}$ son consideradas condicionalmente independientes dados los valores de los parámetros y las variables predictoras. Un supuesto importante en el ajuste del modelo de regresión es el de varianzas iguales, es decir, $var\ (y_i|\theta, X) = \sigma^2$. Enseguida hacemos que $\theta = (\beta_1, \ldots, \beta_k, \sigma^2)$ represente el vector de parámetros desconocidos y asumimos que los errores $\varepsilon_i = y_i - E(y_i|\beta, X)$ son independientes y normalmente distribuidos con media cero y varianza σ^2. La formulación del modelo se completa asumiendo que (β, σ^2) tiene la previa no informativa típica

$$g(\beta, \sigma^2) \propto \frac{1}{\sigma^2} \qquad (8)$$

La función de densidad conjunta de (β, σ^2) se representa como el producto

$$g(\beta, \sigma^2|y) = g(\beta|y, \sigma^2)g(\sigma^2|y) \quad (9)$$

La distribución posterior del vector de regresión β condicional sobre la varianza del error σ^2, $g(\beta|y, \sigma^2)$, es normal multivariada con media $\hat{\beta}$ y matriz de varianzas y covarianzas V_β, donde

$$\hat{\beta} = (X`X)^{-1}X`y, \qquad V_\beta = (X`X)^{-1} \quad (10)$$

De acuerdo a Hamada (2008), es posible generar muestreos de una distribución a posteriori usando una función de verosimilitud adecuada y, por supuesto una correcta distribución a priori para los parámetros. Específicamente Hamada menciona el Metropolis-Hasting y el Muestreo de Gibbs.

El comando R enseguida genera el modelo y sus coeficientes.

```
> fit=lm(RATIO~PROCESSED+ACCEPTED+NP,data=data, x= TRUE,y=TRUE)
> summary(fit)
Call:
lm(formula = RATIO ~ PROCESSED + ACCEPTED + NP, data = data,
   x = TRUE, y = TRUE)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.024424	-0.005786	-0.001225	0.005689	0.024442

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.882e-01	5.417e-02	12.704	1.39e-06 ***
PROCESSED	-1.618e-05	1.145e-06	-14.126	6.13e-07 ***
ACCEPTED	2.495e-05	1.651e-06	15.111	3.64e-07 ***
NP	-7.115e-06	8.406e-06	-0.846	0.422

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 (significancia de los predictores)

Residual standard error: 0.01851 on 8 degrees of freedom
Multiple R-squared: 0.9731, Adjusted R-squared: 0.963
F-statistic: 96.47 on 3 and 8 DF, p-value: 1.274e-06
y el modelo puede escribirse como

$$RATIO = O.6882 - 0.00001618PROCECCED + 0.00002495ACCEPTED - 0.000007115NP \quad (11)$$

El modelo ha sido validado con datos adicionales obteniéndose ajustes similares. Este modelo podrá ser construido con datos de costos, los cuales por problemas propios de los departamentos de contraloría no ha sido posible construirlo.

BIBLIOGRAFIA

1. Albert,Jim. ((2209), Bayesian Computation with R, Second Edition, Springer.
2. Berger, J.O. , (1985) Statistical Decision Theory and Bayesian Analysis. Springer Verlag, New York.
3. Box & Tiao, (1973), Bayesian Inference in Statistical Analysis, Wiley Classics Library. Print ISBN: 9780471574286
4. Hamada, Michael S., Alyson G. Wilson, C. Shane Reese, Harry F. Martz, (2008), Bayesian Reliability, Springer. ISBN 0172-7397
5. Gelman, Andrew, (2004), Prior Distribution Volume 3 pp 1634-1637 in Encyclopedia of Enviromenmetrics ISBN 0471 899976
6. Robert T. Lund, Remanufacturing: The Experience of the United States and Implications for Developing Countries. Word Bank Technical Paper Number 31, December 1984.