Aplicación de la prueba de bondad de ajuste de Kolmogorov-Smirnov para modelos de pérdida.

Karen Contreras Cortés [a], **Bulmaro Juárez Hernández** [a], **José Miguel Hernández López** [a], **Karina Reyes Juárez** [a]
[a] *Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Avenida San Claudio y 18 Sur, Colonia San Manuel, Ciudad Universitaria, C.P. 72570, Puebla, México.*
karen.ctz81@gmail.com, helojmee@gmail.com, juarezh58@gmail.com, katiapadillarodriguez@gmail.com, kary.reyesjz@gmail.com

Resumen
El presente trabajo, consiste en la propuesta de un modelo que ajuste el comportamiento de los montos para los eventos de pérdida del ramo del sector asegurador de Riesgos Hidrometeorológicos. Así como comprobar la validez del modelo propuesto mediante la prueba de bondad de ajuste de Kolmogorov-Smirnov.

Palabras clave: Pérdida, ajuste, Kolmogorov-Smirnov, riesgos hidrometeorológicos.

1. Introducción

Para evitar que una propiedad quede desprotegida ante cualquier evento natural existen los seguros de daños, los cuales contemplan su protección y amparan tanto al edificio como los contenidos.

Los riesgos CAT o riesgos catastróficos son aquellos que se originan por hechos o acontecimientos naturales de caracter extraordinario (baja frecuencia y elevada intensidad) con impacto económico significativo debido a los daños que generan. Usualmente se dividen en 2 grandes riesgos:

- **Movimientos geológicos**. Terremoto, erupción volcánica, tsunami, etc.
- **Hidro-meteorológicos**. Huracán, tormenta, inundaciones, vientos tempestuosos, etc.

1.1. Factores a considerar.

¿Cómo se evalúan este tipo de riesgos? Se toman en cuenta los siguientes factores:

- **Tipos constructivos**
- **Ubicación**
- **Distribución de valores**

Los tipos constructivos generalmente tienen que ver con el material con el que están diseñados los muros, techos y estructuras que componen la propiedad asegurada y son de mucha ayuda para determinar la calidad y resistencia de la construcción.

Es importante conocer la ubicación del riesgo. Una forma de evaluar la exposición de los riesgos es a través de las zonas CRESTAS (Catastrophe Risk Evaluation and Standardizing Target Accumulations), que están definidas para la mayor parte del mundo y México.

Las zonas CRESTA son la base esencial para el análisis de negociación y cartera de reaseguro. Información de la zona CRESTA es utilizada por la mayoría de aseguradoras para evaluar las tasas de seguros que cobran.

En México además se cuenta con la zonificación de la AMIS, Asociación Mexicana de Instituciones de Seguros, que ofrece un catálogo de zonas de terremoto y riesgos Hidrometeorológicos que permite a las compañías de seguros contar con elementos suficientes para una adecuada toma de decisiones para elaborar tarifas y evaluar siniestros.

2. Base de Datos

Para la integración de la base de datos se recopilan las cifras reportadas en el Sistema Estadístico del Sector Asegurador dentro del apartado de Riesgos Hidrometeorológicos, de los años 2008 a 2019 [6].

En dicha sección se puede consultar la información estadística de mercado que reportan las Instituciones de Seguros y de Fianzas de acuerdo a sus operaciones, ramos o tipos de seguro.

Tabla 1. Base de Datos

Fecha	Entidad	Uso	Tipo de Bien	Monto ajustado	Monto VF
2008	Baja California	Casa Habitación	Edificio	4888.5	7741.28419
2008	Baja California	Casa Habitación	Edificio	9699.6667	15360.1056
2008	Baja California	Casa Habitación	Contenidos	75633	119779.185
...	...	...	...	...	...
2019	Nuevo Leon	No valuables...	Contenidos	1719272.82	1719272.82
2019	D.F. (CDMX)	Agua	Edificio	5690380.52	5690380.42
2019	Sinaloa	Vivienda(Crédito Hipotecario)	Edificio	54829.19	54829.19
2019	Sinaloa	Vivienda(Crédito Hipotecario)	Contenidos	1132.96	1132.96

La base de datos está compuesta por tres segmentaciones que se consideran importantes: Entidad, Uso y Tipo de bien. La primera columna que lleva por nombre *Fecha* se refiere al año en que fue obtenida la información, esta columna sirve solamente de apoyo.

Es de interés conocer la fecha en la que se produjo la pérdida ya que para hacer que los montos reportados en diferentes años sean comparables se llevarán todos al año 2019 (Tabla 1).

Para calcular el valor de las pérdidas al año 2019 se consultó la inflación anual acumulada al mes de Diciembre de cada año publicada por Banxico desde el año 2008 al 2018 [5] (Tabla 2).

Tabla 2. Inflación

Fecha	Inflación Anual	Factor de Acumulación
01/12/2008	6.53	1.58
01/12/2009	3.67	1.49
01/12/2010	4.4	1.44
01/12/2011	3.82	1.37
01/12/2012	3.57	1.32
01/12/2013	3.97	1.28
01/12/2014	4.08	1.23
01/12/2015	2.13	1.188
01/12/2016	3.36	1.16
01/12/2017	6.77	1.12
01/12/2018	4.83	1.05

3. Caso de Estudio

De acuerdo a la información del Sistema Estadístico del Sector Asegurador (SESA 2019) que corresponde a Riesgos Hidrometeorológicos [4], el 76.29 % del total de la prima emitida se concentra en solo 11 entidades (Tabla 3).

Por lo que el modelo propuesto tomará una segmentación de la base de datos con respecto a la Entidad y el Tipo de Bien asegurado.

Tabla 3. Porcentaje de prima emitida por entidad (2019)

Entidad	% Prima Emitida	Prima Emitida
Cuidad de México	25.27 %	2,801,114,008
Quintana Roo	15.37 %	1,703,156,690
Baja California Sur	6.07 %	672,959,772
Jalisco	5.11 %	566,772,833
Veracruz	4.59 %	508,768,491
Nuevo León	4.53 %	502,281,809
Tamaulipas	3.99 %	441,798,038
Estado de México	3.83 %	424,730,076
Sinaloa	2.96 %	328,539,855
Yucatán	2.41 %	267,505,270
Guanajuato	2.15 %	238,052,989
Resto	23.71 %	2,628,492,024
Total	100 %	11,084,171,853

Entidad: Distrito Federal (Ciudad de México)

Tipo de Bien: Edificio

De manera que con el uso de esta segmentación se tiene un total de 2,284 observaciones. Más aún, se hace uso de la escala logarítmica en las cifras.

Definición 1. (*Escala logarítmica*) Una *escala logarítmica* es una escala de medida que utiliza el logaritmo de una cantidad física en lugar de la propia cantidad.
La presentación de datos en una escala logarítmica puede ser útil cuando los datos cubren una amplia gama de valores, el logaritmo los reduce a un rango más manejable.

4. Construcción de la Función de Distribución Acumulada Empírica

Gracias a la segmentación de datos realizada, se cuenta con un conjunto de 2,284 observaciones del valor de las pérdidas. Se calcularán algunas descripciones numéricas de interés (Tabla 4).

Tabla 4. Descripciones Numéricas

Descr. Num.	Valor
Mínimo	0.0472
Máximo	20.7509
Rango	20.7037
Q_1	9.2848
Mediana	10.5025
Q_3	11.9118
Rango Intercuantil	2.6270
Media	10.5530
Varianza	5.5842
Desviación Estándar	2.3631
Asimetría	-0.2864
Curtosis	2.8382

Ahora se encuentra la Función de Distribución Empírica para el conjunto de datos y para ello se hace uso de intervalos de clase.

Definición 2. (*Función de Distribución Empírica*) Sean $x_1, \ldots, x_n$ observaciones de una variable cuantitativa. La función de distribución empírica se define como:

$$F(x) = \frac{\#\{x_i : x_i \leq x\}}{n}$$

La Función de Distribución Empírica es una manera alternativa de representar la información completa de un conjunto de datos numéricos. Estima la verdadera Función de Densidad Acumulativa de los puntos en la muestra, y se garantiza virtualmente que converge con la distribución verdadera a medida que el tamaño de la muestra se hace lo suficientemente grande.

Definición 3. (*Intervalo de Clase*) Se denomina *intervalo* o *intervalo de clase* a cada una de las partes en las que se puede subdividir el recorrido de una variable.
Los intervalos pueden ser de la misma o distinta amplitud, es decir, que la diferencia entre sus límites superiores e inferiores puede ser igual o no.

Si bien, por comodidad, se trabajó con intervalos de clase de amplitud constante e idéntica. Asimismo, se conviene que los intervalos de clase sean abiertos por su correspondiente extremo inferior y cerrados por su extremo superior.
Los intervalos de clase se emplean si las variables toman un número grande de valores o si la variable es contínua. Se agrupan los valores en intervalos que tengan la misma amplitud denominados clases, a cada clase se le asigna la frecuencia correspondiente.

Definición 4. (*Regla de Sturges*) La regla de Sturges es un criterio que se aplica para calcular el número de clases o intervalos precisos requeridos para representar gráficamente un conjunto de datos estadísticos en una distribución de frecuencia para datos agrupados.
El número óptimo de intervalos k es dado por la expresión:

$$K = 1 + \frac{log(N)}{log(2)}$$

Donde:
K es el número de clases.
N es el número total de observaciones de la muestra.
log(·) es el logaritmo común de base 10.
El valor de K es común redondearlo al entero más cercano.

Para nuestro caso de estudio, el número de clases está dado por:

$$\begin{aligned} K &= 1 + \frac{log(N)}{log(2)} = 1 + \frac{log(2284)}{log(2)} \\ &= 1 + \frac{3.358696}{0.301029} = 1 + 11.157346 \\ &= 12.157346 \approx 12 \end{aligned}$$

De manera que, el cálculo de la Función de Distribución Empírica puede observarse en la siguiente tabla:

Tabla 5. Función de Distribución Empírica

Intervalo		Frec.	$\hat{f}$(x)	$\hat{F}$(x)
LI	LS			
0	0.0472	3	0.0013	0.0013
0.0472	1.7725	14	0.0061	0.0074
1.7725	3.4978	2	0.0009	0.0083
3.4978	5.2231	23	0.0101	0.0184
5.2231	6.9484	63	0.0276	0.0460
6.9484	8.6737	301	0.1318	0.1778
8.6737	10.3990	687	0.3008	0.4785
10.3990	12.1243	684	0.2995	0.7780
12.1243	13.8497	350	0.1532	0.9313
13.8497	15.5750	109	0.0477	0.9790
15.5750	17.3003	36	0.0158	0.9947
17.3003	19.0256	9	0.0039	0.9987
19.0256	20.7509	3	0.0013	1.000

Nota. LI:=Límite Inferior, LS:=Límite Superior, Frec.:=Frecuencia, $\hat{F}$(x):= Func. distr. empírica.

5. Distribuciones Propuestas

En la rama de seguros son importantes dos variables aleatorias, a saber: el número de siniestros y la cuantía por reclamación. El comportamiento de la primera dará idea de la frecuencia con la que se producen dichos siniestros. La segunda, ayuda a valorar en términos económicos la ocurrencia de los mismos.

Para el caso de estudio se utiliza la variable aleatoria de cuantía por reclamación, en este caso de carácter continuo, pues modeliza una variable económica medida en dinero, que como la variable tiempo, se considera continua.

Las distribuciones comunmente empleadas, por tener generalmente un buen ajuste a los datos históricos son la distribución Pareto, Gamma o Weibull para la cuantía de los siniestros [3].

5.1. Distribución Pareto. .

Es lo que se conoce en estadística como una distribución de colas anchas, es decir, su densidad refleja una característica que tambien corrobora la varianza de esta distribución, hay probabilidad alta de que la variable tome valores alejados de la media. Sirve para modelizar siniestros con grandes costes o cuantías por reclamación.

Definición 5. (*Distribución Pareto*) Se dice que la v.a. X tiene distribución de Pareto de parámetros k, α, si su soporte es el subconjunto de números reales $[k,\infty)$ y su función de densidad viene dada por

$$f(x) = \begin{cases} \frac{\alpha k^{\alpha}}{x^{\alpha+1}}, & k < x, \\ 0, & k \geq x \end{cases}$$

Normalmente k es un parámetro fijado de antemano, a partir de los datos iniciales de la aplicación. El parámetro α es un parámetro de ajuste, cuyo valor se estima o determina a posteriori.

Su función de distribución es

$$F(x) = \ 1 - \frac{k}{x}^{\alpha}, \quad x > k$$

además, su valor esperado y varianza son

$$E(X) = \frac{\alpha k}{\alpha - 1}, \quad \alpha > 1,$$

y

$$V(X) = \frac{\alpha k^2}{(\alpha-2)(\alpha-1)^2}, \ \ \alpha > 2.$$

Se ha visto que la existencia de los dos primeros momentos depende del valor del parámetro α. Es por ello, que a menudo conviene estimar su valor, a partir de los datos, para tener una mayor información del comportamiento de la variable.

5.2. Distribución Gamma. .

Definición 6. (*Distribución Gamma*) Se dice que la v.a. X sigue una distribución Gamma de parámetros $\alpha > 0$, $\beta > 0$, y se denota por $\Gamma(\alpha,\beta)$, si su función de densidad viene dada por

$$f(x) = \begin{cases} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leq 0, \end{cases}$$

donde, $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$.

Su función de distribución es

$$F(x) = \ \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \beta x), \quad x > 0,$$

donde $\gamma(\alpha,\beta x)$ representa la gamma incompleta evaluada en α y βx. Además, su valor esperado y varianza son

$$E(X) = \frac{\alpha}{\beta} \ \text{ y } \ V(x) = \frac{\alpha}{\beta^2}.$$

5.3. Distribución de Weibull. .

Esta distribución se usa habitualmente para modelar la compensación de las reclamaciones de los trabajadores. Proporciona una adecuada representación de la distribución de la pérdida.

Definición 7. (*Distribución de Weibull*) Se dice que la v.a. X sigue una distribución de Weibull de parámetros $\alpha > 0$, $\beta > 0$ y se denota por $W(\alpha,\beta)$, si su función de densidad viene dada por

$$f(x) = \begin{cases} \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

La función de distribución es

$$F(x) = 1 - e^{\left(-\frac{x}{\beta}\right)^{\alpha}}.$$

además, su valor esperado y varianza son

$$E(x) = \beta\Gamma(\frac{1}{\alpha}+1), \ \text{ y } \ V(x) = \beta^2(\Gamma(\frac{2}{\alpha}+1) - \Gamma^2(\frac{1}{\beta}+1)).$$

6. Modelo Propuesto

En seguida se evaluan las funciones de distribución propuestas bajo los intervalos de clase previamente calculados (Tabla 6).

Es importante mencionar que por el momento, los parámetros son tomados arbitrariamente como:
$\alpha = 2$ (parámetro 1) y $\beta = 1$ (parámetro 2) para las tres distribuciones propuestas ya que tales parámetros serán ajustados más adelante.

Una vez obtenidos los valores para las distribuciones propuestas con parámetros arbitrarios, se obtuvo para cada una de estas distribuciones el error cuadrático medio con respecto a la Función de Densidad Empírica. (Tabla 7)

Y con ayuda de la herramienta *Solver* se obtienen aquellos valores de los parámetros α y β para los cuales el error cuadrático medio es mínimo (Tabla 8).

Tabla 6. Distribuciones Propuestas

Intervalo		Gamma	Weibull	Pareto
LI	LS			
0	0.0472	0.0233	0.0233	0.0230
0.0472	1.7725	0.5878	0.5878	0.4698
1.7725	3.4978	0.8260	0.8260	0.6362
3.4978	5.2231	0.9266	0.9266	0.7231
5.2231	6.9484	0.9690	0.9690	0.7765
6.9484	8.6737	0.9869	0.9869	0.8126
8.6737	10.3990	0.9945	0.9945	0.8387
10.3990	12.1243	0.9977	0.9977	0.8584
12.1243	13.8497	0.9990	0.9990	0.8738
13.8497	15.5750	0.9996	0.9996	0.8862
15.5750	17.3003	0.9998	0.9998	0.8964
17.3003	19.0256	0.9999	0.9999	0.9049
19.0256	20.7509	1.0000	1.0000	0.9121

Nota. LI:= Límite Inferior, LS:= Límite Superior

Tabla 7. ECM con Parámetros Arbitrarios

	Gamma	Weibull	Pareto
	0.0005	0.0005	0.0005
	0.3368	0.3368	0.2138
	0.6687	0.6687	0.3943
	0.8248	0.8248	0.4966
	0.8520	0.8520	0.5337
	0.6547	0.6547	0.4031
	0.2662	0.2662	0.1297
	0.0482	0.0482	0.0065
	0.0046	0.0046	0.0033
	0.0004	0.0004	0.0086
	0.0000	0.0000	0.0097
	0.0000	0.0000	0.0088
	0.0000	0.0000	0.0077
Suma ECM	3.6570	3.6570	2.2162

En seguida se obtienen los valores para las distribuciones propuestas con los nuevos parámetros (Tabla 9).

Así, se puede observar que la distribución que presenta el menor ECM respecto a la Función Empírica es la distribución Gamma(25.13 , 0.42). De manera que se desea probar:

$$H_0 : \hat{F}(x) = F(x), \quad \text{con } \alpha = 0.05,$$

donde, $\hat{F}(x) :=$ Función de Distribución Empírica.
$F(x) := FDA$ Gamma$(25.13, 0.42)$.

7. Validez del Modelo Propuesto

Una prueba de bondad de ajuste cuando $F_0(x)$ es continua, es la basada en la estadística de Kolmogorov-Smirnov.

Tabla 8. Parámetros por Solver

Parámetros	Gamma	Weibull	Pareto
α	25.13	5.71	42 730 258.44
β	0.42	11.32	518 033 154.78

Tabla 9. ECM con Parámetros dados por Solver

	Gamma	Weibull	Pareto
	0.0000	0.0000	0.0000
	0.0001	0.0001	0.0165
	0.0001	0.0001	0.0587
	0.0003	0.0000	0.1100
	0.0003	0.0002	0.1523
	0.0000	0.0003	0.1111
	0.0000	0.0003	0.0095
	0.0000	0.0000	0.0213
	0.0000	0.0007	0.0627
	0.0000	0.0004	0.0654
	0.0000	0.0000	0.0551
	0.0000	0.0000	0.0428
	0.0000	0.0000	0.0326
Suma ECM	0.00089	0.002	0.74

Esta se basa en una comparación entre las funciones de distribución acumulativa que se observan en la muestra ordenada y la distribución propuesta bajo la hipótesis nula. Si esta comparación revela una diferencia suficientemente grande entre las funciones de distribución muestral y propuesta, entonces la hipótesis nula de que la distribución es $F_0(x)$ se rechaza.

Considérese la hipótesis nula como:

$$H_0 : F(x) = F_0(x),$$

en donde, $F_0(x)$ se especifica en forma completa.
Si $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, denotan a las observaciones ordenadas de una muestra de tamaño n y defínase la función de distribución acumulativa muestral, la cual es equivalente a la función de distribución empírica y puede darse en términos de las observaciones ordenadas [1].

$$S_n(x) = \begin{cases} 0, & si \quad x < x_{(1)}, \\ \vdots & \\ k/n, & si \quad x_{(k)} \leq x \leq x_{(k+1)}, \\ \vdots & \\ 1, & si \quad x \geq x_n. \end{cases}$$

En otras palabras, para cualquier valor ordenado x de la muestra aleatoria, $S_n(x)$ es la proporción del número de valores en la muestra que son iguales o menores a x. Ya que $F_0(x)$ se encuentra completamente especificada, es posible evaluar a $F_0(x)$ para algún valor deseado de x, y entonces comparar este último con el valor correspondiente de $S_n(x)$. Si la hipótesis nula es verdadera, entonces es lógico esperar que la diferencia sea relativamente pequeña. La estadística de Kolmogorov-Smirnov se define como:

$$D_n = máx \mid S_n(x) - F_0(x) \mid .$$

La estadística D_n, tiene una distribución que es independiente del modelo propuesto bajo la hipótesis nula. Por esta razón, se dice D_n, es una estadística independiente de la distribución. Lo anterior da como resultado que la función de distribución de D_n pueda evaluarse sólo en función del tamaño de la muestra y después usarse para cualquier $F_0(x)$.

Para un tamaño α del error de tipo I, la región crítica es de la forma:

$$P\left(D_n > \frac{c}{\sqrt{n}}\right) = \alpha.$$

De acuerdo con lo anterior, la hipótesis H_0, se rechaza si para algún valor x observado el valor de D_n, se encuentra dentro de la región crítica de tamaño α.

Se comprobó la validez del modelo propuesto mediante la prueba de bondad de ajuste de Kolmogorov-Smirnov expuesta anteriormente.

Recuérdese que se desea contrastar o probar:

$$H_0 : \hat{F}(x) = F(x).$$

Donde:

$\hat{F}(x)$:= Función de Distribución Empírica.

$F(x)$:= FDA Gamma(25.13 , 0.42).

Con un nivel de significancia $\alpha = 0.05$.

Para el cálculo de la estadística D_n se tiene:
Entonces:

$$\begin{aligned} D_n &= máx \mid F_E(x) - F_T(x) \mid \\ &= 0.0184. \end{aligned}$$

De acuerdo a la tabla de valores cuantiles superiores de la distribución de la estadística D_n de Kolmogorov-Smirnov [2], la región crítica es de la forma:

$$P\left(D_n > \frac{1.36}{\sqrt{2284}}\right) = 0.05.$$

Tabla 10. Diferencias

Intervalo LI	Intervalo LS	$\hat{F}$(x)	F(x)	$\mid\hat{F}(x) - F(x)\mid$
0	0.0472	0.0013	0.0000	0.0013
0.0472	1.7725	0.0074	0.0000	0.0074
1.7725	3.4978	0.0083	0.0000	0.0083
3.4978	5.2231	0.0184	0.0009	0.0175
5.2231	6.9484	0.0460	0.0276	0.0184
6.9484	8.6737	0.1778	0.1805	0.0028
8.6737	10.3990	0.4785	0.4851	0.0065
10.3990	12.1243	0.7780	0.7722	0.0059
12.1243	13.8497	0.9313	0.9277	0.0036
13.8497	15.5750	0.9790	0.9828	0.0038
15.5750	17.3003	0.9947	0.9968	0.0021
17.3003	19.0256	0.9987	0.9995	0.0008
19.0256	20.7509	1.0000	0.9999	0.0001

Nota. LI:=Límite Inferior, LS:=Límite Superior

Donde:
Si $D_n > \frac{1.36}{\sqrt{2284}}$ se Rechaza H_0.
Si $D_n < \frac{1.36}{\sqrt{2284}}$ No se rechaza H_0.

Finalmente, como $D_n = 0.0184 < 0.0285 = \frac{1.36}{\sqrt{2284}}$.

No se rechaza H_0 y se concluye, con un nivel de significancia de 0.05, en base a los datos observados, que dichos datos provienen de una distribución $Gamma(25.13, 0.42)$.

8. Resultados

La prueba de bondad de Ajuste de Kolmogorov-Smirnov nos dá las bases teóricas necesarias para poder concluir que los montos de pérdida del sector asegurador en el ramo de Riesgos Hidrometeorológicos ocurridos en el Distrito Federal por edificios se ajustan a una distribución Gamma(25.13 , 0.42) con un nivel de confianza del 95 %.

Se puede observar gráficamente que en efecto, la distribución $Gamma(25.13, 0.42)$ propuesta es un buen ajuste para los datos observados ver Figura 1.

Finalmente, interesa usar esta información para analizar ciertos eventos, tales como:

- La probabilidad que el monto de pérdida sea mayor a \$500,000.00. Esto es,

$$\begin{aligned} P(X > \ln(500000)) &= P(X > 13.122363) \\ &= 1 - P(X \leq 13.122363) \\ &= 1 - F(13.122363) \\ &= 1 - 0.878165 \\ &= 0.121834. \end{aligned}$$

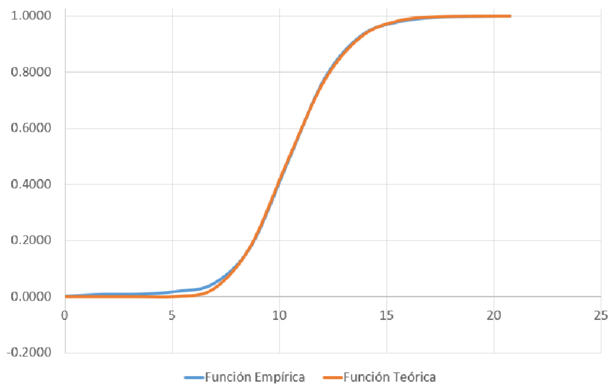

Figura 1. Distribución Empírica vs Distribución Teórica

Por lo tanto, el monto de pérdida será mayor a \$500,000.00 con probabilidad 0.121834.

- La probabilidad que el monto de pérdida esté entre \$250,000.00 y \$500,000.00.

$$\begin{aligned} P(\ln(250000) \leq X \leq \ln(500000)) &= P(X \leq \ln(500000)) \\ &\quad - P(X \leq \ln(250000)). \end{aligned}$$

Calculando cada término:

$$\begin{aligned} P(X \leq \ln(500000)) &= P(X \leq 13.122363) \\ &= F(13.122363) \\ &= 0.878165. \end{aligned}$$

$$\begin{aligned} P(X \leq \ln(250000)) &= P(X \leq 12.4292) \\ &= F(12.4292) \\ &= 0.809638. \end{aligned}$$

Así, se obtiene:

$$\begin{aligned} P(\ln(250000) \leq X \leq \ln(500000)) &= P(X \leq \ln(500000)) \\ &\quad - P(X \leq \ln(250000)) \\ &= 0.878165 - 0.809638 \\ &= 0.068527. \end{aligned}$$

Esto es, el monto de pérdida estará entre \$250,000.00 y \$500,000.00 con probabilidad 0.068527.

Agradecimientos

Agradecemos a los profesores de la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla por compartir su conocimiento y experiencias con todos nosotros.

En especial a los profesores tanto del cuerpo de Actuaría como de Probabilidad y Estadística, ya que gracias a ellos hemos conseguido unificar en éste trabajo los conocimientos adquiridos a lo largo de nuestra formación universitaria.

Referencias

[1] C. Canavos, G. *Probabilidad y Estadística. Aplicaciones y métodos.*, Edo. de México: McGraw-Hill/Interamericana de México, 1988.

[2] F.J Massey, Jr. *The Kolmogorov-Smirnov test for goodness of fit.*, J. Amer Statistical Assoc. 46 (1951), 68-78.

[3] Garin, M. & Gonzalez, M. *Modelos estocasticos de estadistica actuarial*, Universidad del Pais Vasco/ Euskal Herriko Unibertsitatea, España.

[4] Asociación Mexicana de Instituciones de Seguros *Sistema Estadístico del Sector Asegurador del Ramo de Riesgos Hidrometeorológicos de daños SESA 2019.*, México, D.F, 2020.

[5] Banco de México *Sistema de Información Económica.*, https://www.banxico.org.mx/tipcamb/main.do?page=inf&idioma=sp , 2021.

[6] Comisión Nacional de Seguros y Fianzas *Información estadística detallada del sector asegurador.Riesgos Hidrometeorológicos.*, https://www.cnsf.gob.mx/EntidadesSupervisadas/InstitucionesSociedadesMutualistas/Paginas/RiesgosHidrometeorologicos.aspx, 2020.