

Ajuste de Funciones de Distribución Probabilística

Profesor Efraín Domínguez



Facultad de Estudios Ambientales y Rurales
Departamento de Ecología y Territorio
e.dominguez@javeriana.edu.co

7 de septiembre de 2015

Contenido

- 1 Función de Distribución Empírica F_*
 - Probabilidad de Excedencia/No Excedencia
 - Función de Distribución Empírica F_*
- 2 Función Teórica de Distribución Probabilística
 - Propiedades de las Funciones de Distribución
 - Curvas Teóricas de Densidad Probabilística
- 3 Verificación de Bondad de Ajuste
 - Criterios de Bondad de Ajuste
 - Análisis de Resultados del Ajuste
- 4 Análisis de Implementación en Python
- 5 Bibliografía

Probabilidad de No Excedencia

Definición

La función de distribución de probabilidades $-F(x)-$ de una magnitud aleatoria X representa la probabilidad de que la magnitud aleatoria X no exceda un valor determinado x . De otro modo se escribe cómo:

$$F(X) = p(X \leq x). \quad (1)$$

Probabilidad de Excedencia

Definición

En algunos casos es más conveniente tratar a la función de distribución de probabilidades $-F(x)-$ de una magnitud aleatoria X como la probabilidad de que la magnitud aleatoria X exceda un valor determinado x . Algunos autores, para diferenciar representan esta función de distribución de probabilidades de excedencia como $P(x)$, diferenciandola de la función de distribución de probabilidades de no excedencia $F(x)$, y describiendola como:

$$P(X) = p(X \geq x). \quad (2)$$



Formulas Prob. de Exced./No Exced. Empírica F_*

Número	Autor	alfa	Formula
1)	Weibull;Kritskiy-Menkel	0.00	$F(Q_i) = \frac{i}{(n+1)}$
2)	Gringorton	0.44	$F(Q_i) = \frac{(i-0,44)}{(n+0,12)}$
3)	Hazen	0.55	$F(Q_i) = \frac{(i-0,5)}{(n)}$
4)	Chegodaeiev	0.30	$F(Q_i) = \frac{(i-0,3)}{(n+0,4)}$
5)	Blókhinov	0.40	$F(Q_i) = \frac{(i-0,4)}{(n+0,2)}$
6)	Distribución Gumbel	0.44	$F(Q_i) = \frac{(i-0,44)}{(n+0,12)}$
7)	Distribución normal	3.0/8.0	$F(Q_i) = \frac{(i-\frac{3,0}{8,0})}{(n+\frac{1,0}{4,0})}$

Cuadro : Fuente: [Druzhinin, Sikan, 2001]

Mediana de la Distribución

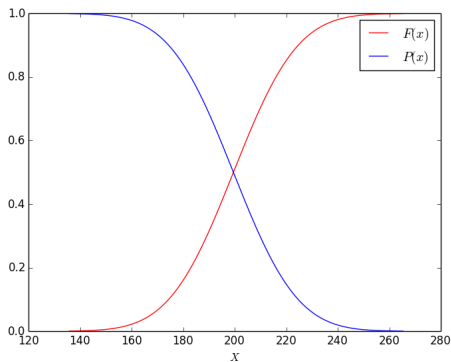


Figura : Probabilidad de Excedencia / No Excedencia

Propiedades de las Funciones de Distribución

$$\lim_{x \rightarrow -\infty} F(x) = 0; \quad \lim_{x \rightarrow -\infty} P(x) = 1; \quad (3)$$

$$\lim_{x \rightarrow +\infty} F(x) = 1; \quad \lim_{x \rightarrow +\infty} P(x) = 0; \quad (4)$$

$$F(x) \geq 0 \quad P(x) \geq 0 \quad \forall x \quad (5)$$

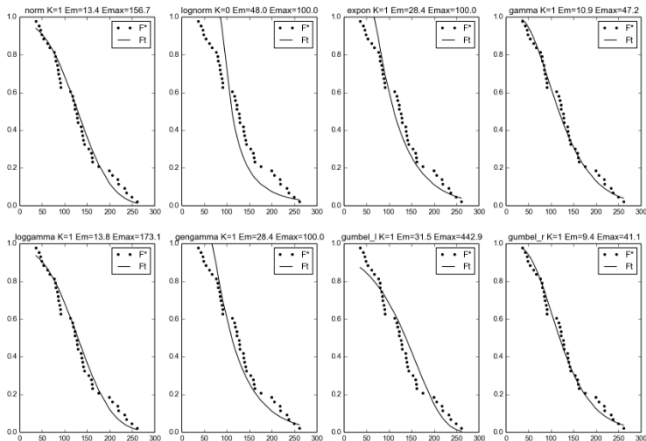
$$F(X_2) \geq F(X_1), \text{ si } x_2 > x_1; \quad P(X_2) \leq P(X_1), \text{ si } x_2 < x_1 \quad (6)$$

Curvas Teóricas de Densidad Probabilística

Distribución	Formula
Normal	$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$
Log Normal	$p(x) = \frac{e^{-\frac{(\ln(x)-\ln(\bar{x}))^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}x}$
Exponencial	$p(x) = \lambda e^{-\lambda x} \forall x \geq 0$
Gamma - γ	$p(x) = \frac{\lambda^2 x^{(a-1)} e^{-\lambda x}}{\Gamma(a)}$
Weibull min	$p(x) = cx^{(c-1)} e^{-(x^c)}$
Weibull max	$p(x) = c(-x)^{(c-1)} e^{-(-x)^c}$
Gumbel L	$p(x) = e^{(x+e^x)}$
Gumbel R	$p(x) = e^{-(x+e^{(-x)})}$

Cuadro : Distribuciones Teóricas

Ejemplo de Inspección Visual



Errores Promedio y Máximo de Ajuste

Error Absoluto

$$\epsilon_i = 100 * \left| \frac{F_{*i} - F(x)_i}{F_{*i}} \right| \quad (7)$$

Error Promedio

$$\bar{\epsilon} = 100 \frac{\sum \left| \frac{F_{*i} - F(x)_i}{F_{*i}} \right|}{n} \quad (8)$$

Error Máximo

$$\epsilon_{max} = \max \left(100 * \left| \frac{F_{*i} - F(x)_i}{F_{*i}} \right| \right) \quad (9)$$

Criterio de Kolmogorov

Postulación de Hipótesis Binaria

H_0 - La distribución teórica analizada representa bien la distribución empírica (concuera con esta).

H^1 - La distribución teórica analizada no representa bien la distribución empírica (no concuera con esta).

Divergencia máxima de ajuste - D

$$D = \max_{i=1}^n |F_{*i} - F(x)_i| \quad (10)$$

λ **Empírico**

$$\lambda = D\sqrt{n} \quad (11)$$

Estadístico λ_α Teórico

Distribución de Kolmogorov $p(D\sqrt{n} \geq \lambda)$

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2\lambda^2} \quad (12)$$

α	0.40	0.30	0.20	0.10	0.05
λ_α	0.89	0.97	1.07	1.22	1.36
α	0.025	0.01	0.005	0.001	0.0005
λ_α	1.48	1.63	1.73	1.95	2.03

Cuadro : λ_α Teórico

¿Cómo decidir?

Verificación de Hipotesis:

Si se cumple que $\lambda \leq \lambda_\alpha \Rightarrow$ la Hipotesis nula H_0 no se rechaza.

Verificación de Errores Estadísticos

- Errores promedio $> 20\% \Rightarrow$ Ajuste Aceptable.
- Errores promedio $\leq 20\% \Rightarrow$ Ajuste Aceptable.
- Errores promedio $\leq 15\% \Rightarrow$ Ajuste Bueno.
- Errores promedio $\leq 10\% \Rightarrow$ Ajuste Excelente.

¿Cómo implementar este análisis en Python?

El código comentado para el ajuste de funciones de distribución se puede descargar en:

- 1 Código para el ajuste de una función Gamma;
- 2 Código para el ajuste de varias funciones teóricas de distribución probabilística.

Bibliografía



Andrey Kolmogorov (1931)

On analytical methods in probability theory

Math. Ann. 104, 415 – 458.



C. Haan Thomas (1977)

Statistical Methods in Hydrology

Iowa, Iowa State University Press, 378.



V. Druzhinin, A. Sikan, (2001)

Statistic methods for the analysis of hydrometeorological information

Saint Petersburg, RGGMU, 168.