Distribuciones para valores extremos

Teoría de valores extremos

1. La Teoría de Valores Extremos (Extreme value theory) consiste en el empleo de una serie de técnicas estadísticas para la identificación y modelado de observaciones extremas u outliers.

2. Su objeto es determinar que tan extrema puede ser la mayor o menor observación registrada de un fenómeno aleatorio, es decir, estudia el comportamiento del valor *máximo o mínimo* de una variable aleatoria.

3. El comportamiento inusual de una variable aleatoria merece una consideración especial, ya que puede tener un gran impacto para las decisiones que se desprendan del análisis de la información a la que pertenece.

4. Para explicar este tipo de sucesos que ocurren, que ocurren, generalmente, con muy baja frecuencia, pero que tienen influencia muy significativa sobre todo un modelo. La *Teoría de Valores Extremos* emplea métodos matemáticos basados en comportamientos asintóticos, distribuciones, procesos estocásticos y leyes limite.

5. Diferentes investigaciones provenientes de múltiples disciplinas científicas, han desarrollado métodos para cuantificar eventos extremos y sus consecuencias de un modo estadísticamente óptimo, dando lugar a unas distribuciones de probabilidad que permiten la modelación de los valores máximos o mínimos de una variable aleatoria.

   De forma simplificada, nuestro problema es el siguiente:

   Dada una muestra independiente $X_1, X_2, ..., X_n$ de una distribución desconocida, $\mathbf{F}$, queremos *estimar la cola* de $\mathbf{F}$. Los problemas más importantes son:

   - Las observaciones en la cola de la distribución son escasas.

   - Por lo general, queremos estimar valores por encima del valor máximo de la muestra.

   - Las técnicas usuales de estimación de densidades ajustan bien en las zonas donde los datos tienen mayor densidad, pero pueden ser inadecuadas para estimar las colas.

- Los modelos correspondientes a esta teoría de valores extremos, tienen aplicaciones en muchas áreas, una de las principales es: las ciencias ambientales, donde se estudian valores extremos, por ejemplo en: Nivel de una presa, velocidad del viento, nivel de un río, concentración de contaminantes, niveles de precipitación pluvial, etc...

- No obstante , nosotros nos enfocamos en aplicarla dentro del marco del seguro. En esta área, el análisis de siniestralidad extrema es de gran interés, puesto que constituye un riesgo que pone en peligro la estabilidad y solvencia de entidades aseguradoras.

Recordemos una de las definiciones de convergencia, que aprendieron en los cursos de probabilidad básica:

Definición: Convergencia en distribución

La sucesión de variables aleatorias $X_1, X_2, ...$ converge en distribución a X, si para todo punto x en donde la función $F_X(x)$ es continua, se cumple que:

$$\lim_{n\to\infty} F_{X_n}(x) \quad = \quad F_X(x)$$

En este curso nosotros vamos a trabajar únicamente el caso univariado.

Corolario

Sea $\{X_i\}_{i=1}^n$ una muestra de v.a.i.i.d. Definimos $M_n = máx\{x_i\} \quad \forall i$. Entonces para cualquier distribución que tengan los $X_i's$, si existe $a_n > 0$ y $b_n \in \mathbb{R}$, *(constantes de normalización)* tales que:

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{M_n - b_n}{a_n} \leq z\right] \quad = \quad G(z)$$

Donde $G(z)$ es una función de distribución (acumulada), llamando M a la v.a cuya función de distribución es G, decimos entonces que:

$$\frac{M_n - b_n}{a_n} \quad \xrightarrow{d} \quad M$$

Entonces, en caso de encontrar dichas a_n y b_n donde $\frac{M_n - b_n}{a_n}$ tenga una convergencia en distribución a M, entonces la función de distribución (acumulada) de M será G, que basado en el teorema de Fisher-Tippett-Gnedenko, solo podrá tener una de las siguientes formas:

Distribuciones de valores extremos

$M \sim Weibull(\alpha, \beta, \lambda > 0)$

$$G(z) = \begin{cases} exp\left[-\left(-\left(\frac{z-\beta}{\alpha}\right)\right)^{\lambda}\right] & \text{si } z < \beta \\ 1 & \text{si } \underbrace{z \geq \beta}_{\text{Soporte}} \end{cases}$$

$M \sim Gumbel(\alpha, \beta)$

$$G(z) \quad = \quad exp\left[-exp\left(-\left(\frac{z-\beta}{\alpha}\right)\right)\right] \quad \underbrace{\forall z \in \mathbb{R}}_{\text{Soporte}}$$

$M \sim Fréchet(\alpha, \beta, \lambda > 0)$

$$G(z) = \begin{cases} 0 & \text{si } z \leq \beta \\ exp\left[-\left(\frac{z-\beta}{\alpha}\right)^{-\lambda}\right] & \text{si } \underbrace{z > \beta}_{\text{Soporte}} \end{cases}$$

Observación: * n es el tamaño de muestra ($\{X_i\}_{i=1}^n$).

Si $\{X_i\}_{i=1}^n$ son una muestra de v.a.i.i.d entonces la función de distribución de $M_n = máx\{X_i\}$ $\forall i$, será:

$$\begin{aligned} F_{M_n}(t) &= \mathbb{P}[M_n \leq t] \\ &= \mathbb{P}[X_1 \leq t, X_2 \leq t, ..., X_n \leq t] \\ &\overset{indep}{=} \prod_{i=1}^{n} \mathbb{P}[X_i \leq t] \\ &\overset{i.d}{=} \mathbb{P}^n[X \leq t] \\ &= F_X^n(t) \end{aligned}$$

Por lo tanto

$$F_{M_n}(t) = F_X^n(t) \quad \forall t \in \mathbb{R}$$

Entonces:

$$\begin{aligned} \mathbb{P}\left[\frac{M_n - b_n}{a_n} \leq z\right] &= \mathbb{P}\left[M_n \leq a_n z + b_n\right] \\ &= F_{M_n}(a_n z + b_n) \\ &= F_X^n(a_n z + b_n) \end{aligned}$$

Por lo tanto:

$$\mathbb{P}\left[\frac{M_n - b_n}{a_n} \leq z\right] = F_X^n(a_n z + b_n)$$

Noten que los limites a los que se llegan son de la forma e^x, entonces es considerable aprender el siguiente limite:

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{n} = e^{x} \qquad \forall x\in\mathbb{R} \quad (x \text{ NO depende de n})$$

Observen que el soporte de todas las distribuciones mencionadas son diferentes, por lo que observar detenidamente el soporte al calcular el limite puede ser una guía importante.

Ejemplo:

Sea $X \sim exp(1)$.

$$F_X(x) = (1-e^{-x})\mathbb{I}_{(x\geq 0)}(x)$$

Entonces:

$$\begin{aligned} F_{M_n}(x) &= F_X^n(x) \\ &= (1-e^{-x})^n\mathbb{I}_{(x\geq 0)}(x) \end{aligned}$$

Ahora, considerando las constantes de normalización $a_n = 1$ y $b_n = ln(n)$ entonces:

$$\begin{aligned} \lim_{n\to\infty}\mathbb{P}\left[\frac{M_n-b_n}{a_n}\leq z\right] &= \lim_{n\to\infty} F_X^n(a_nz+b_n) \\ &= \lim_{n\to\infty}(1-e^{-(a_nz+b_n)})^n\mathbb{I}_{((a_nz+b_n)\geq 0)}(a_nz+b_n) \\ &= \lim_{n\to\infty}(1-e^{-(ln(n))})^n\mathbb{I}_{\underbrace{((z+ln(n))\geq 0)}_{\text{Observación}}}(z+ln(n)) \\ &= \lim_{n\to\infty}(1-e^{-z}\cdot e^{ln(1/n)})^n \\ &= \lim_{n\to\infty}(1+\frac{-e^{-z}}{n})^n \\ &= e^{-e^{-z}} \end{aligned}$$

Observación:

$$z + ln(n) \geq 0 \Leftrightarrow z \geq -ln(n) \xrightarrow[n\to\infty]{} z \geq -\infty$$

$$\therefore \quad z \in \mathbb{R}$$

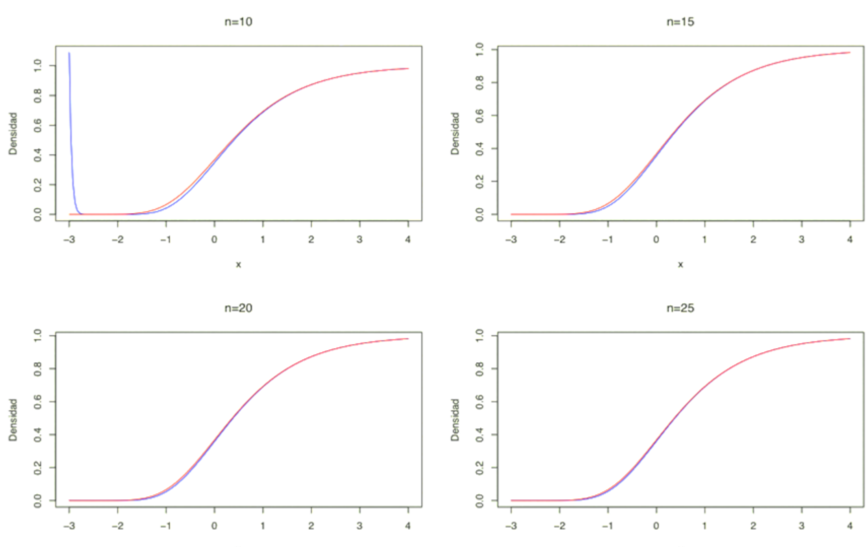
Lo que nos da como resultado:

$$\lim_{n\to\infty} \mathbb{P}[\tfrac{M_n - b_n}{a_n} \leq z] = \lim_{n\to\infty} (1 - e^{-(z+In(n))})^n \mathbb{I}_{(z+In(n)\geq 0)}(z + In(n)) = e^{-e^{-z}} \forall z \in \mathbb{R}$$

Gracias al soporte identificamos que $M \sim Gumbel$,luego buscamos sus parámetros, pero una vez ya identificada es fácil notar que si $\alpha = 1, \beta = 0$:

$$\Rightarrow G(x) = exp\{-exp\{-(\tfrac{x-\beta}{\alpha})\}\} = e^{-e^{-(\frac{x-0}{1})}} = e^{-e^{-x}} \forall z \in \mathbb{R}$$

$$\therefore \tfrac{M_n - b_n}{a_n} = M_n - ln(n) \xrightarrow[n\to\infty]{d} M \sim Gumbel(\alpha = 1, \beta = 0)$$



Resultado gráfico de valor extremo Gumbel

Ejemplo 2: Consideremos $X \sim Unif(0,1) \Longrightarrow \mathrm{F}_X(x) = x \overset{\star}{\mathbb{I}}_{(0,1)}(x)$

Notación: $\overset{\star}{\underset{(a,b)}{\mathbb{I}}}(x) = \begin{cases} si\ x \leq a \Rightarrow\ F_X(x) \equiv 0 \\ si\ x \in\ (a,b) \Rightarrow \overset{\star}{\mathbb{I}}_{(a,b)}(x) = 1 \\ si\ x \geq\ b \Rightarrow\ F_X(x) \equiv 1 \end{cases}$

Tomando una muestra $\{\mathrm{X}_i\}_{i=1}^{n}, entonces:$

$$F_{M_n}\ (\mathrm{t}) = \mathbb{P}\ [\ M_n = \underset{\forall_i}{máx}\{\mathrm{X}_i\} \leq t] = \mathrm{F}_X^n(\mathrm{t}) = t^n\ \overset{\star}{\mathbb{I}}_{(0,1)}(t)$$

Pequeña observación:

$$F_{M_n}\ (\mathrm{t}) = t^n\ \overset{\star}{\mathbb{I}}_{(0,1)}(t) \Rightarrow \tfrac{\partial}{\partial t}F_{M_n}(t) = \mathrm{f}_{M_n}\ (\mathrm{t}) = nt^{n-1}\ \mathbb{I}_{(0,1)}(t)\ \dashrightarrow \text{Indicadora usual}$$

$$\Rightarrow f_{M_n}(t) = nt^{n-1}\mathbb{I}_{(0,1)}(t) = \tfrac{\Gamma(n+1)}{\Gamma(n)\Gamma(1)}t^{n-1}(1-t)^{1-1}\mathbb{I}_{(0,1)}(t)$$

$$\therefore \mathrm{M}_n = \underset{\forall_i}{máx}\{\mathrm{X}_i\} \sim Beta(n,1)\ (si\ \mathrm{X}_i\ \perp'\ s\ Unif(0,1))$$

$$\mathbb{E}[\ M_n] = \tfrac{n}{n+1};\ \mathbb{V}\mathrm{ar}\ (\ M_n\ ) = \tfrac{n(1)}{(n+1+1)(n+1)^2} = \tfrac{n}{(n+2)(n+1)^2}$$

Consideremos $\mathrm{b}_n = 1\ y\ a_n = \frac{1}{n}, entonces:$

$$\lim_{n\to\infty} \mathbb{P}\left[\tfrac{M_n - b_n}{a_n} \leq t\right] = \lim_{n\to\infty} F_{M_n}(a_n t + b_n) = \lim_{n\to\infty} {F_x}^n(a_n t + b_n) = \lim_{n\to\infty}[a_n t + b_n]^n\ \overset{\star}{\underset{(0,1)}{\mathbb{I}}}(a_n t + b_n)$$
$$= \lim_{n\to\infty}\ \underbrace{\left[\frac{t}{n}+1\right]^n}_{\textit{Atendiendo el límite}}\ \underbrace{\overset{\star}{\mathbb{I}}_{(0,1)}\left(\frac{t}{n}+1\right)}_{\textit{Atendiendo la indicadora}}$$

Observación

- Atendiendo el límite

$\lim_{n\to\infty}(1+\frac{t}{n})^n = e^t$

- Atendiendo la indicadora

$0 \leq 1 + \frac{t}{n} \leq 1 \Leftrightarrow -1 \leq \frac{t}{n} \leq 0 \Leftrightarrow -n \leq t \leq 0$
$Cuando$ n$\to \infty$ $entonces:$

$$t \in (-\infty, 0)$$

Por lo tanto:
$\lim\limits_{n\to\infty} \mathbb{P}[\frac{M_n-b_n}{a_n} \leq t] = \lim\limits_{n\to\infty} (1+\frac{t}{n})^n \overset{\star}{\mathbb{I}}_{(0,1)}(1+\frac{t}{n}) = e^t \overset{\star}{\mathbb{I}}_{(-\infty,0)}(t)$

$$\Rightarrow G(t) = \begin{cases} e^t & si \ x < 0 \\ 1 & si \ x \geq 0 \end{cases} \to M \sim Weibull$$

Una vez más, gracias al soporte podemos identificar la distribución de valor extremo a la que converge en distribución $\frac{M_n-b_n}{a_n}$. Procedemos a buscar a cuál en específico.
Recordemos que la f.d.a. de una Weibull es:

$$G(z) = \begin{cases} exp[-(-(\frac{z-\beta}{\alpha}))^\lambda] & \text{si} \quad z < \beta \\ 1 & \text{si} \quad z \geq \beta \end{cases}$$

Tomando $\alpha = 1, \beta = 0, \lambda = 1$:

$$G(z) = \begin{cases} exp[-(-(z))] = e^z & \text{si} \quad z < 0 \\ 1 & \text{si} \quad z \geq 0 \end{cases}$$

$\therefore \frac{M_n-b_n}{a_n} \xrightarrow{d} M \ \sim Weibull(\alpha = 1, \beta = 0, \lambda = 1)$

¡Más aún! Tomando G'(z) $= \underbrace{g(z) = e^z \mathbb{I}_{(-\infty,0)}(z)}_{\text{Función de densidad de M}}$

$$\begin{aligned} \Rightarrow \mathbb{E}[M] &= \int_{-\infty}^{0} z e^z dz \underset{\substack{\downarrow \\ du=-dz \\ u=-z}}{=} \int_{\infty}^{0} (-u)e^{-u}(-du) = \int_{\infty}^{0} u e^{-u} du \\ &= -\int_{0}^{\infty} u e^{-u} du = -\int_{0}^{\infty} u f_{Exp(1)}(u) du = -\mathbb{E}[Exp(1)] = -1 \\ \therefore \mathbb{E}[M] &= -1 \end{aligned}$$

La siguiente pregunta es ¿Estará ligado al valor anterior con la convergencia que tiene $\frac{M_n - b_n}{a_n}$? Tomando:

$$\begin{aligned}
\mathbb{E}\left[\frac{M_n - b_n}{a_n}\right] &= \frac{\mathbb{E}[M_n] - b_n}{a_n} \quad \text{(pues } a_n \text{ y } b_n \text{ no son v.a.'s)} \\
&= \frac{(\frac{n}{n+1}) - 1}{\frac{1}{n}} = n\left[\frac{n - n - 1}{n+1}\right] = -\frac{n}{n+1} = -\left(\frac{n+1-1}{n+1}\right) \\
&= -\left(1 - \frac{1}{n+1}\right) = \frac{1}{n+1} - 1 \\
\therefore \quad & \mathbb{E}\left[\frac{M_n - b_n}{a_n}\right] = \frac{1}{n+1} - 1
\end{aligned}$$

Pero luego:

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{M_n - b_n}{a_n}\right] = -1 = \mathbb{E}[M]$$

Noten que no necesitábamos conocer la distribución de M para lograr obtener su esperanza.

Queda como ejercicio para el lector calcular $Var(M)$ y comprobar que

$$\lim_{n\to\infty} Var(\frac{M_n - b_n}{a_n}) = Var(M) \text{ en este caso}$$

VaR & T-VaR

VaR

El VaR **es un valor** utilizado para **cuantificar el riesgo**. En términos formales, el VaR mide la máxima pérdida esperada en un intervalo de tiempo determinado, bajo condiciones normales del mercado y bajo un nivel de confianza dado. Traducido a una forma sencilla y como se estará manejando desde el punto de vista de Teoría del Riesgo y la estadística, el VaR diremos que es un cuantil; dada una variable aleatoria X que mida el monto de pérdida de una compañía de seguros el VaR a cierto nivel de confianza p se define como:

$$VaR_p(X) := F_X^{-1}(p) = \pi_p \longleftrightarrow F_X(VaR_p(X)) = p = F_X(\pi_p)$$

T-VaR

El valor de la cola en riesgo (TVaR), es una medida de riesgo asociada con el valor más general en riesgo. Cuantifica el **valor esperado** de la pérdida dado que se ha producido un evento fuera de un nivel de probabilidad dado. Dada una variable aleatoria X que mide el monto de pérdida de una compañía de seguros y una cierta cantidad π_p definimos el TVaR como:

$$TVaR_p(X) := \mathbb{E}[X|X > \pi_p] = \frac{\int_{\pi_p}^{\infty} x f_X(x)dx}{1-F(\pi_p)}$$

$$TVaR_p(X) \overset{P.D.}{=} \pi_p + e(\pi_p) = \frac{\int_{\pi_p}^{\infty} x f_X(x)dx}{1-F(\pi_p)}, \text{ con: } e(\pi_p) = \frac{\int_{\pi_p}^{\infty}(x-\pi_p) f_X(x)dx}{1-p}$$

Demostración

Tenemos que:

$$\pi_p + \frac{\int_{\pi_p}^{\infty} x f_X(x)dx - \int_{\pi_p}^{\infty} \pi_p f_X(x)dx}{1-p} = \frac{\pi_p(1-p-(1-p)) + \int_{\pi_p}^{\infty} x f_X(x)dx}{1-p} = \frac{\int_{\pi_p}^{\infty} x f_X(x)dx}{1-p} = TVaR_p(X)$$

Queda como ejercicio para el lector que:

$$TVaR_p(X) = \frac{\int_p^1 VaR_u(X)du}{1-p}$$

	$f_X(x)$ (Densidad)	π_p (VaR)	TVaR
Exponencial $x \sim exp(\lambda)$ con $\lambda > 0$	$\lambda e^{-\lambda x}\mathbb{I}_{(0,\infty)}^{(x)}$	$-\frac{ln(1-p)}{\lambda}$	$\frac{1-ln(1-p)}{\lambda}$
Pareto 1 $x \sim Pareto1(\alpha, \theta)$, con $\alpha > 0, \theta > 0$	$\frac{\alpha\theta^\alpha}{x^{\alpha+1}}\mathbb{I}_{(\theta,\infty)}^{(x)}$	$\theta(1-p)^{\frac{1}{\alpha}}$	$\frac{\alpha\theta(1-0)^{-\frac{1}{\alpha}}}{\alpha-1}$
Normal $x \sim Normal(\mu, \sigma^2)$ $\mu \in \mathbb{R}, \sigma > 0$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}\mathbb{I}_{-\infty,\infty}^{(x)}$	$\mu + \sigma\Phi^{-1}(p)$	$\mu + \frac{\sigma}{1-p}\phi(\Phi^{-1}(p))$
Log-Normal $x \sim$ $Log-Normal(\mu, \sigma^2)$, $\mu \in \mathbb{R}, \sigma > 0$	$\frac{1}{x\sqrt{2\pi\sigma^2}}e^{\frac{-(ln(x)-\mu)^2}{2\sigma^2}}\mathbb{I}_{0,\infty}^{(x)}$	$e^{\sigma\Phi^{-1}(p)+\mu}$	$\left(\frac{e^{\mu+\frac{1}{2}\sigma^2}}{1-p}\right)\left(1-\Phi\left(\frac{ln(\pi_p)-\mu-\sigma^2}{\sigma}\right)\right)$

Aclaración:

Sea $Y = X|X > a$ con X una v.a continua. Si $t > a$, entonces:

$$\Rightarrow F_Y(t) \doteqdot \mathbb{P}[Y \leq t] = \mathbb{P}[X \leq t|X > a] = \frac{\mathbb{P}[a < X \leq t]}{\mathbb{P}[X > a]} = \frac{F_X(t) - F_X(a)}{S_X(a)} \quad \forall t \geq a$$

Obs: si $t \leq a \Rightarrow \mathbb{P}[X \leq t|X > a] = 0$

$$\Rightarrow f_Y(t) = \frac{d}{dt}F_Y(t) = \frac{f_X(t)}{S_X(a)}$$

$$\boxed{\therefore f_{X|X>a}(t) = \frac{f_X(t)}{S_X(a)}\mathbb{I}_{t>a}(t)}$$

De tal manera que:

$$TVaR \doteqdot \mathbb{E}[X|X > \pi_p] = \int_{\pi_p}^{\infty} tf_{X|X>\pi_p}(t)dt = \int_{\pi_p}^{\infty} t\frac{f_X(t)}{S_X(\pi_p)}dt = \frac{\int_{\pi_p}^{\infty} tf_X(t)dt}{S_X(\pi_p)}$$

De ahí obtenemos el cálculo para el TVaR. Como ya hemos visto, hay maneras alternativas para calcular este valor. Anteriormente se mostró que:

$$\boxed{TVaR_p(X) = \pi_p + \frac{\int_{\pi_p}^{\infty}(x-\pi_p)f_X(x)dx}{1-p}}$$

A partir de este último resultado veamos otro par de detalles.
Recordemos lo siguiente:

PROPOSICIÓN. (DESIGUALDAD DE MARKOV)

Sea $X \geq 0$ una variable aleatoria con esperanza finita. para cualquier $\varepsilon > 0$,

$$P(X > \varepsilon) \geq \frac{E(X)}{\varepsilon}$$

Demostración

$$\begin{aligned} E(X) &= E(X1_{(X\geq\varepsilon)} + X1_{(X<\varepsilon)}) \\ &\geq E(X1_{(X\geq\varepsilon)}) \\ &\geq E(\varepsilon 1_{(X\geq\varepsilon)}) \\ &= \varepsilon P(X \geq \varepsilon) \quad \square \end{aligned}$$

Dicho de otra manera, si X es continua con función de supervivencia $S_X(t)$ entonces:

$$0 \leq tS_X(t) \leq \mathbb{E}[X] \quad \forall t \geq 0$$

Entonces $\lim_{t\to\infty} tS_X(t)$ existe y es no negativo si $\mathbb{E}[X] < \infty$
Más aún:

$$0 \leq \lim_{t\to\infty} tS_X(t) = \lim_{t\to\infty} t\int_t^\infty f_X(x)dx = \lim_{t\to\infty}\int_t^\infty tf_X(x)dx \leq \lim_{t\to\infty}\int_t^\infty \underset{\downarrow}{\overset{\text{pues } t\leq x\leq\infty}{}} xf_X(x)dx = 0$$

$0 \leq \int_0^\infty xf_X(x)dx = \mathbb{E}[X] < \infty$ e integrar asi la reduce $\hookleftarrow$

$$\boxed{\therefore \lim_{t\to 0} tS_X(t) = 0 \quad \text{si} \quad X \geq 0 \quad \text{y} \quad \mathbb{E}[X] < \infty}$$

Una identidad interesante a notar es, dados los supuestos anteriores:

$$\boxed{\int_{\pi_p}^\infty (x-\pi_p)f_X(x)dx = \int_{\pi_p}^\infty S_X(x)dx}$$

Dem

$$\begin{aligned}
\underbrace{\int_{\pi_p}^b (x-\pi_p)f_X(x)dx}_{\substack{u=(x-\pi_p)\Rightarrow du=dx \\ dv=f_X(x)dx\Rightarrow v=F_X(x)}} &= (x-\pi_p)F_X(x)|_{\pi_p}^b - \int_{\pi_p}^b F_X(x)dx = (b-\pi_p)F_X(b) - \int_{\pi_p}^b (1-S_X(x)dx) \\
&= (b-\pi_p)F_X(b) - (b-\pi_p) + \int_{\pi_p}^b S_X(x)dx \\
&= (b-\pi_p)(F_X(b)-1) + \int_{\pi_p}^b S_X(x)dx = (\pi_p - b)S_X(b) + \int_{\pi_p}^b S_X(x)dx \\
&= \pi_p S(b) - bS_X(b) + \int_{\pi_p}^b S_X(x)dx
\end{aligned}$$

$$\therefore \lim_{b\to\infty}\int_{\pi_p}^b (x-\pi_p)f_X(x)dx = \lim_{b\to\infty}\int_{\pi_p}^b S_X(x)dx \quad \square$$

Habiendo probado esto, se deduce que si X es v.a no-negativa con esperanza finita, entonces:

$$\boxed{TVaR_p(X) = \pi_p + \frac{\int_{\pi_p}^\infty S_X(t)dt}{1-p}}$$

Observación:

- $\phi(t) \doteqdot$ función de densidad de una normal estándar (evaluada en t).

- $\Phi(t) \doteq$ función de distribución de una normal estándar (evaluada en t).

La notación $VaR_p(X)$ puede cambiar por VaR_p si ya se entiende sobre qué se calcula. Igual para $TVaR_p$

¿De dónde sale la tabla anterior?

Ejemplo 1: $X \sim Exp(\lambda)$ obtengamos $VaR_p(X)$

Sea $VaR_p(X) = \pi_p$ entonces:

$$\begin{aligned}
\mathbb{P}[X \leq \pi_p] &= p = 1 - e^{\lambda \pi_p} \\
&\Leftrightarrow e^{-\lambda \pi_p} = 1 - p \Leftrightarrow \lambda \pi_p = -ln(1-p) \\
&\Leftrightarrow \pi_p = -\frac{ln(1-p)}{\lambda} = VaR_p \qquad \text{Lo que teníamos: } -\frac{ln(1-p)}{\lambda}
\end{aligned}$$

Obtengamos $TVaR_p(x)$

Primero:

$$\begin{aligned}
\int_{\pi_p}^{\infty} S_X(x)dx &= \int_{\pi_p}^{\infty} e^{-\lambda x}dx = \frac{1}{\lambda}\int_{\pi_p}^{\infty} \lambda e^{-\lambda x}dx \\
&= \frac{1}{\lambda}\int_{\pi_p}^{\infty} f_X(x)dx = \frac{1}{\lambda}S_X(\pi p) \\
&= \frac{e^{-\lambda \pi_p}}{\lambda} \\
\Rightarrow TVaR_p &= \pi_p + \frac{\int_{\pi_p}^{\infty} S_X(x)dx}{S_X(\pi_p)} = -\frac{ln(1-p)}{\lambda} + \frac{\frac{\cancel{e^{-\lambda(\pi_p)}}}{\lambda}}{\cancel{e^{-\lambda(\pi_p)}}} \\
&= \underbrace{-\frac{ln(1-p)}{\lambda}}_{} + \underbrace{\frac{1}{\lambda}}_{} = \frac{1 - ln(1-p)}{\lambda} \\
\mathbb{E}[X|X > \pi_p] &= \qquad \pi_p \qquad + \mathbb{E}[X]
\end{aligned}$$

Pérdida de memoria de la la distribución exponencial.

Lo que teníamos: $\frac{1-ln(1-p)}{\lambda}$

Ejemplo 2: Sea $X \sim Pareto1(\alpha, \theta)$ Obtengamos $VaR_p(x)$

$$\begin{aligned}
F_X(x) &= 1 - \left(\frac{\theta}{x}\right)^{\alpha} \\
&\Rightarrow F_X(\pi_p) \doteq \mathbb{P}[X \leq \pi_p] = p = 1 - \left(\frac{\theta}{\pi_p}\right)^{\alpha} \\
&\Leftrightarrow \left(\frac{\theta}{\pi_p}\right)^{\alpha} = 1 - p \Leftrightarrow \frac{\theta}{\pi_p} = (1-p)^{-\frac{1}{\alpha}} \\
&\Leftrightarrow \pi_p = \frac{\theta}{(1-p)^{\frac{1}{\alpha}}} = \theta(1-p)^{-\frac{1}{\alpha}}
\end{aligned}$$

Lo que teníamos: $\theta(1-p)^{-\frac{1}{\alpha}}$

Queda como ejercicio para el lector obtener VaR_p de una v.a $X \sim log - Normal(\mu, \sigma^2)$.
Hint: $Y \sim N(\mu, \sigma^2) \Leftrightarrow X = e^Y \sim log - Normal(\mu, \sigma^2)$
De tal manera que:

$$F_X(x) = F_Y(ln(x)) = \Phi\left(\frac{ln(x) - \mu}{\sigma^2}\right)$$

Script:"Valores extremos y VaR"