MAGNA EDUCATIO AD MELIOR

UNIVERSIDAD DE SANTANDER

Universidad de Santander
UDES

VIGILADA MINEDUCACIÓN | SNIES 2832

Estadística

Variable aleatoria continua

Daniel Martínez Bello

Universidad de Santander
Maestría en Biotecnología

Octubre de 2025

Variables aleatorias

FDA, función de sobrevida y cuantiles

Valor esperado

- Variable aleatoria continua

Reglas sobre los valores esperados

Varianza

- Distribuciones continuas

Variable aleatoria

- Una **variable aleatoria** es el resultado numérico de un experimento aleatorio.
- Una variable aleatoria corresponde a todos los posibles resultados del experimento junto con la probabilidad de cada resultado.
- Se encuentran variables aleatorias **discretas** o **continuas**.
- La variable aleatoria continua puede tomar cualquier valor en la recta real o en un subconjunto de la recta real.
  - $P(X \in A)$

Función de densidad de probabilidad

La función de densidad de probabilidad (FDP) es una función asociada a una variable aleatoria continua

> *El aréa bajo la FPD corresponde a las probabilidades de esa variable aleatoria*

Una FDP f valida debe satisfacer,

1. $f(x) \geq 0$ para todo x
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Ejemplo

Suponiendo que el tiempo en años desde el momento del diagnóstico a la muerte de las personas de cierta condición de salud

$$f(x) = \begin{cases} \frac{e^{-x/8}}{8} & \text{para } x > 0 \\ 0 & \text{otro valor} \end{cases}$$

Escrito en otra forma: $f(x) = \frac{1}{8}e^{-x/8}$ para $x > 0$.

Esto es una densidad valida?

1. e elevado a cualquier valor es positivo
2. $$\int_0^\infty f(x)dx = \int_0^\infty e^{-x/8}/8dx = -e^{-x/8}\Big|_0^\infty = 1$$

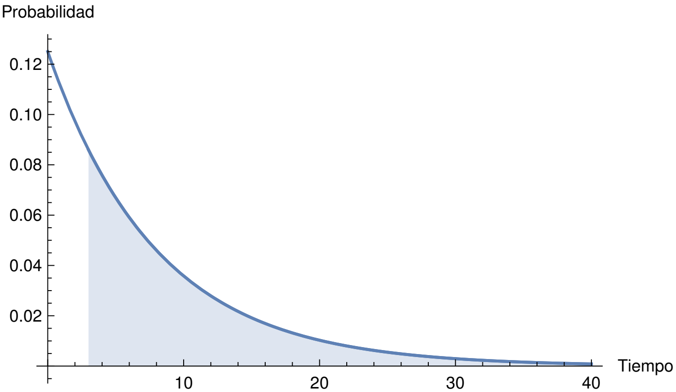
Ejemplo
Cuál es la probabilidad que una persona seleccionada de esta distribución sobreviva mas de 3 años?

$$P(X \geq 3) = \int_3^{\infty} \frac{e^{-t/8}}{8} dt = -e^{-t/8}\Big|_3^{\infty} = e^{-3/8} \approx 0.687$$

Aproximación en R

```
pexp(3, 1/8, lower.tail = FALSE)
```

Ejemplo



Función de distribución acumulada y función de sobrevida

- La **función de distribución acumulada** (FDA) de una variables aleatoria X se define como la función

$$F(x) = P(X \leq x)$$

- Esta definición se aplica independiente de si X es discreta o continua
- La **función de sobrevida** de una variable aleatoria X se define como

$$S(x) = P(X > x)$$
$$S(x) = 1 - F(x)$$

- Para las variables aleatorias continuas, la FDP es la derivada de la FDA

$$f(x) = \frac{d}{dx}F(x)$$

Propiedades de la FDA

- $0 \leq F(x) \leq 1$, para todo número real x.
- F es monótona creciente.
- $\text{Lím}_{x \to \infty} F(x) = 1$ y $\text{Lím}_{x \to -\infty} F(x) = 0$
- $P(a < X \leq b) = F(b) - F(a)$
- F es continua por la derecha.

Ejemplo: cuál es la función de sobrevida y la FDA de la densidad exponencial considerada antes?

$$S(x) = \int_x^\infty \frac{e^{-t/8}}{8} dt = -e^{-t/8}\Big|_x^\infty = e^{-x/8}$$

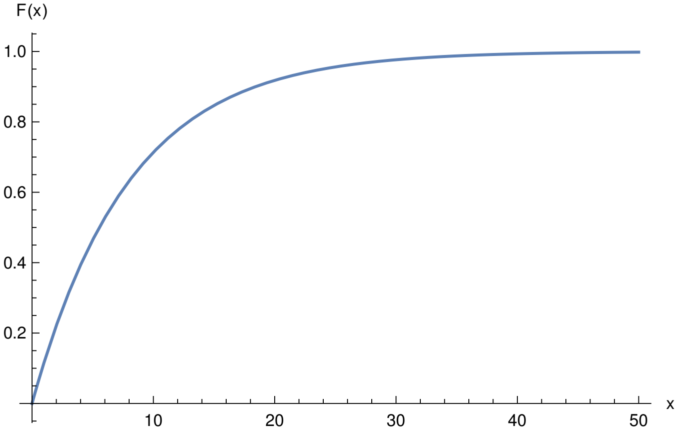
Luego

$$F(x) = 1 - S(x) = 1 - e^{-x/8}$$

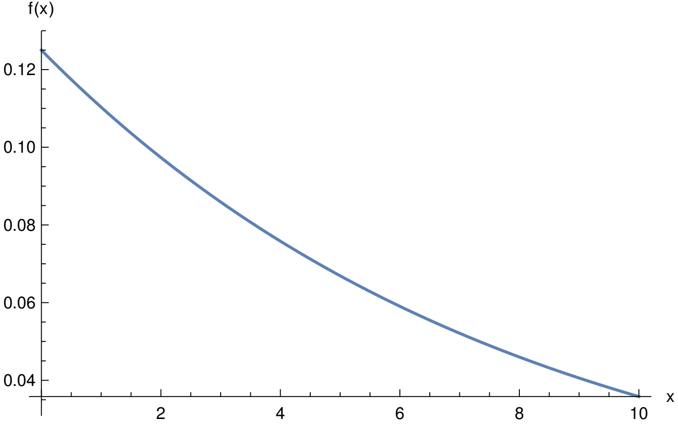
Se puede obtener la FDP a partir de la FDA

$$f(x) = F'(x) = \frac{d}{dx}(1 - e^{-x/8}) = \frac{1}{8}e^{-x/8}$$

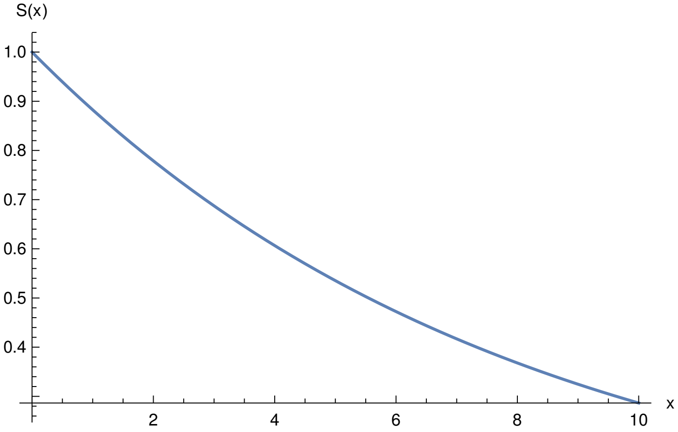
FDA



FDP



Sobrevida



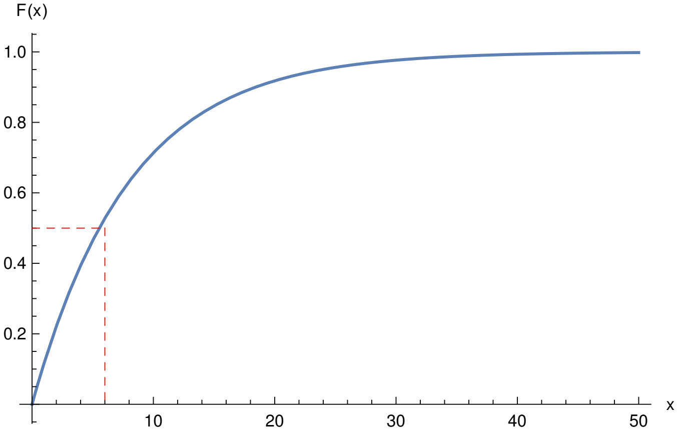
Cuantiles

- El α^{esimo} **cuantil** de una distribución con función de distribución F es el punto x_α de tal forma que

$$F(x_\alpha) = \alpha$$

- Un **percentil** es un cuantil con α expresado como porcentaje
- La **mediana** es el 50^{esimo} percentil

Mediana (Percentil 50)



Ejemplo

- Cual es el 25^{esimo} percentil de la distribución de sobrevida exponencial
- Se resuelve (para x)

$$
\begin{aligned}
.25 &= F(x) \\
&= 1 - e^{-x/8}
\end{aligned}
$$

que da la solución $x = -\log(.75) \times 8 \approx 2.30$

- Por esto, 25% de los sujetos en la población viven menos de 2.30 años
- R aproxima cuantiles

```
qexp(.25, 1/8)
```

Modelos de probabilidad

- Los modelos de probabilidad conectan los datos a la población a traves de unos supuestos.
- El investigador asume un modelo de probabilidad, y genera inferencia, es decir, a partir de muestras al azar, se obtiene información de la población.

Variable aleatoria continua

- Para una variable aleatoria continua, X, con densidad, f, el valor esperado es

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Ejemplo

- Considere la FDP donde $f(x) = 1$ para x entre cero y uno
- (Es una FDP valida?)
- Suponga que X sigue esta densidad; cuál es su valor esperado?

$$E[X] = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = 1/2$$

Reglas sobre los valores esperados

- El valor esperado es un operador lineal
- Si a y b no son aleatorios X y Y son dos variables aleatorias, entonces
  - $E[aX + b] = aE[X] + b$
  - $E[X + Y] = E[X] + E[Y]$
- *En general* si g es una función que no es lineal

$$E[g(X)] \neq g(E[X])$$

- Por ejemplo, $E[X^2] \neq E[X]^2$

La varianza

- La varianza de una variable aleatoria es una medida de *dispersión*
- Si X es una variable aleatoria con media μ, la varianza de X se define como

$$\text{Var}(X) = E[(X-\mu)^2]$$

Que corresponde al cuadrado de la distancia de X a la media

- Densidades con varianzas grandes estan mas dispersas que densidades con un varianza baja
- Otra forma de expresarlo

$$\text{Var}(X) = E[X^2] - E[X]^2$$

- Si a es una constante, $\text{Var}(aX) = a^2\text{Var}(X)$
- La raíz cuadrada de la varianza se llama la **desviación estándar**
- La desviación estándar tiene las mismas unidades que X

Distribucion uniforme

Una variable aleatoria continua tiene una distribución uniforme sobre el intervalo $[a, b]$ si los intervalos de posibles valores que tienen igual longitud tienen igual probabilidad de contener un valor observado de X.

$$X \sim \text{Uniforme en el intervalo } [a, b]$$

$$f(x) = \frac{1}{b-a} \text{ para } a \leq x \leq b$$

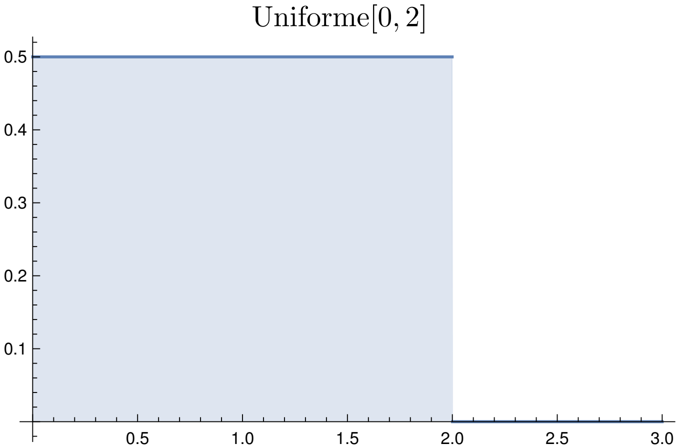
$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{2}$$

Distribución uniforme

La distribución Uniforme se conoce como distribución rectangular. La distribución $U[0, 1]$ se utiliza en métodos de simulación de Monte Carlo.

Distribución uniforme



Distribución Normal

$$X \sim \text{Normal}(\mu, \sigma^2)$$

$$f(x) = \text{Normal}(X|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \text{exp}\left(\frac{1}{2\sigma^2}(X-\mu)^2\right)$$

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

$$moda(X) = \mu$$

Distribución Normal

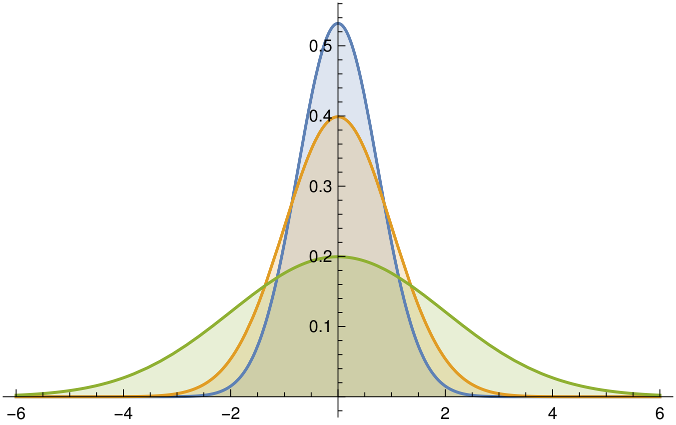
- La situación especial, con $\mu = 0$ y $\sigma = 1$, se le da el nombre de distribución Normal estándar (típica), y este proceso de azar se denota como Z, por tanto $Z \sim \text{Normal}(0; 1)$
- Cualquier variable aleatoria normal $X \sim \text{Normal}(\mu, \sigma^2)$ puede volverse una distribución Normal estándar a través de

$$\frac{X - E(X)}{DE(X)} \sim \text{Normal}(0, 1)$$

- Probabilidades de la forma $P(X \leq x)$ [también denotado como $\Phi(x)$] se encuentran tabuladas. $\Phi(x)$ es la distribución acumulada de una variable aleatoria Normal estándar.

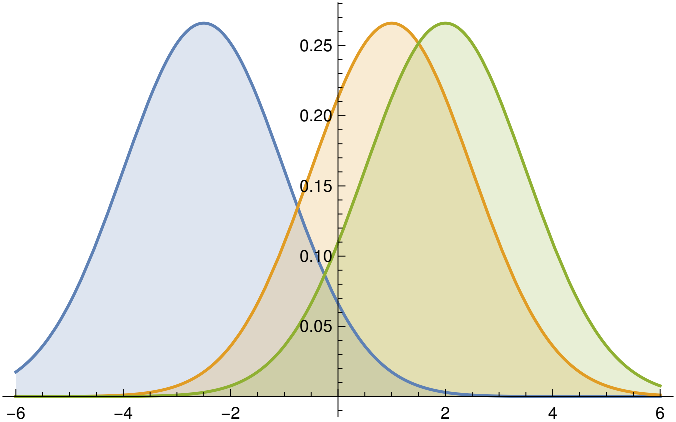
Distribución Normal

Normal[0, 0.75], Normal[0, 1] y Normal[0, 2]



Distribución Normal

Normal$[-2.5, 1.5]$, Normal$[1, 1.5]$ y Normal$[2, 1.5]$



Funciones Gamma y Beta

La función gamma se define como

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

Para enteros no negativos, la función gamma interpola al factorial de x ($x!$), esto es $\Gamma(x+1) = x!$

La función beta se relaciona a la función gamma

$$\text{Beta}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha) + \Gamma(\beta)}$$

Distribución Gamma

$$X \sim \text{Gamma}(\alpha, \beta)$$

$$\text{Forma } \alpha > 0$$

$$\text{Escala inversa } \beta > 0$$

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

$$E(X) = \frac{\alpha}{\beta}$$

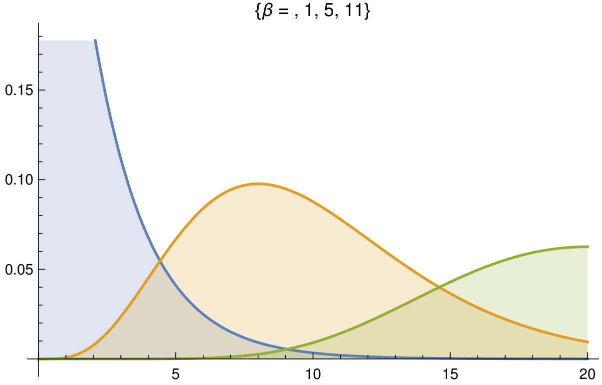
$$Var(X) = \frac{\alpha}{\beta^2}$$

$$moda(X) = \frac{\alpha - 1}{\beta}$$

$$\varphi_X(t) = \left(1 - \frac{it}{\beta}\right)^{\alpha}$$

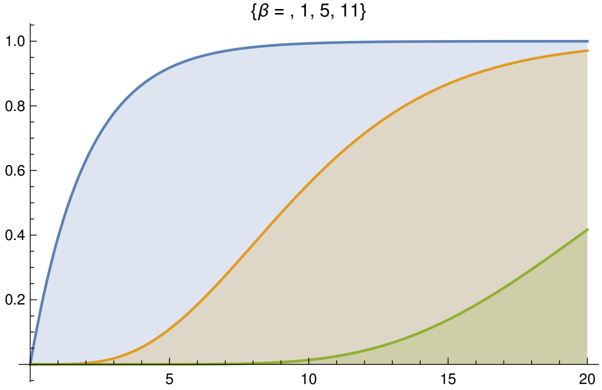
Distribución Gamma

Función de densidad



Distribución Gamma

Función de distribución acumulada



Distribución Gamma

- Modelización de variables con valores no negativos.
- Tiempos de vida de componentes.
- Sean $X_1, \ldots, X_k$, k variables aleatorias independientes tales que X_j tenga como mecanismo generador una distritución Gamma(α, β), para $j = 1, \ldots, k$, entonces la variable $Y = X_1 + \cdots + X_k$ sigue una distribución Gamma$(\alpha_1 + \cdots + \alpha_k, \beta)$. Se demuestra utilizando la función característica

$$\varphi_Y(t) = \prod_{j=1}^{k} \varphi_X(t) = \prod_{j=1}^{k} \left(1 - \frac{it}{\beta}\right)^{\alpha_j}$$

$$= \left(1 - \frac{it}{\beta}\right)^{-\sum_{j=1}^{k} \alpha_j}$$

Distribución Gamma

$$X \sim \text{Gamma}(\alpha, \frac{1}{\beta})$$

$$\text{Forma } \alpha > 0$$

$$\text{Escala } \frac{1}{\beta} > 0$$

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

$$E(X) = \alpha\beta$$

$$Var(X) = \alpha\beta^2$$

$$moda(X) = (\alpha - 1)\beta$$

$$\varphi_X(t) = (1 - it\beta)^\alpha$$

Distribución Gamma Inversa

$$X \sim \text{Gamma-Inv}(\alpha, \beta)$$

$$\text{Forma } \alpha > 0$$

$$\text{Escala } \beta > 0$$

$$f(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{-(\alpha+1)} e^{-\beta/x}$$

$$E(X) = \frac{\beta}{\alpha - 1}$$

$$Var(X) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$$

$$moda(X) = \frac{\beta}{\alpha + 1}$$

Distribución exponencial

La distribución Exponencial sirve para modelar el tiempo al evento de interés que es observado.

$$
\begin{aligned}
X \sim & \operatorname{Exp}(\beta), \text{escala } \beta > 0 \\
f(x) = & \beta e^{-\beta x} \\
E(X) = & \frac{1}{\beta} \\
Var(X) = & \frac{1}{\beta^2} \\
moda(X) = & 0 \\
\text{Gamma}(\alpha = & 1, \beta)
\end{aligned}
$$

Distribución exponencial

- La distribución Exponencial tiene la propiedad de "no tener memoria"

$$P(X > x_0 + x | X > x_0) = \frac{P(X > x_0 + x)}{P(X > x_0)}$$
$$= \frac{e^{-\beta(x_0+x)}}{e^{-\beta x_0}} = P(X > x)$$

- Simulación de una variable aleatoria Exponencial
  - Dadas $X \sim \text{Uniforme}[0,1]$ y $Y \sim \text{Exp}(1)$, la función de distribución acumulada $F_Y(y) = 1 - e^{-y}$.
  - Solucionar $1 - e^{-y} = x$ para y, produce $y = -\log(1-x)$.
  - Si un generador de números pseudoaleatorios produce $x = 0.75$, entonces $y = -\log(0.25) = 1.386$

Distribución Beta

$$X \sim \text{Beta}(\alpha, \beta)$$

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & X \in [0,1] \\ 0 & \text{en otro caso} \end{cases}$$

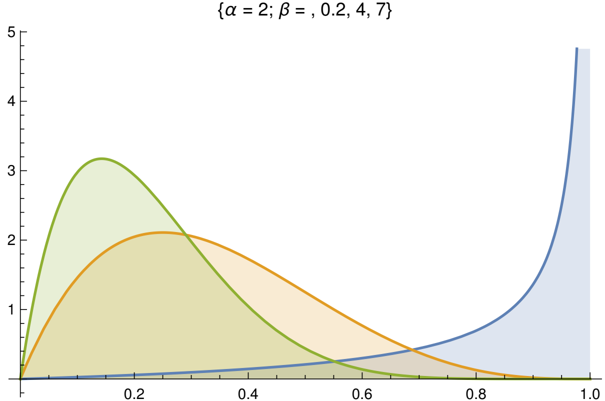
$$E(X) = \frac{\alpha}{\alpha+\beta}$$

$$Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$Moda(X) = \frac{\alpha-1}{\alpha+\beta-2}$$

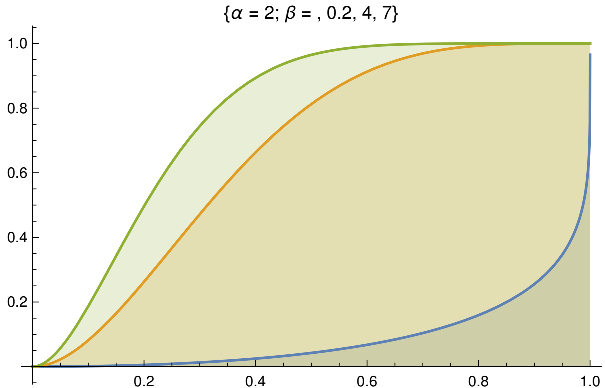
Distribución Beta

Función de densidad



Distribución Beta

Función de distribución acumulada



Distribución Cauchy

$$X \sim \text{Cauchy}(\alpha)(\alpha \in \mathbb{R})$$

$$f(x) = \frac{1}{\pi} \frac{1}{(1 + (x - \alpha)^2)} x \in \mathbb{R}$$

$$E(X) \text{ no existe}$$

$$Var(X) \text{ no existe}$$

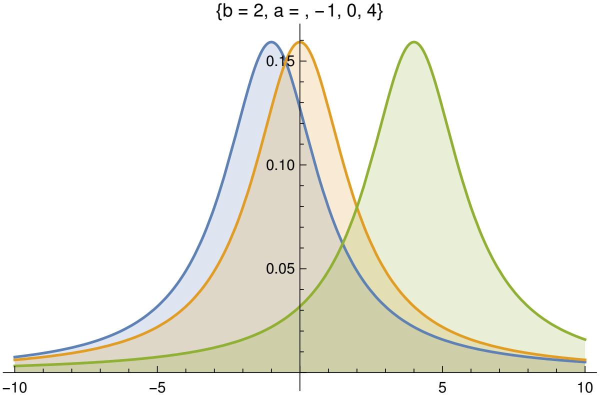
$$\varphi(t) = e^{i\alpha t - |t|}$$

Distribución Cauchy

$$f(x) = \frac{1}{\pi} \frac{1}{b\left(\frac{(x-\alpha)^2}{b^2} + 1\right)}$$

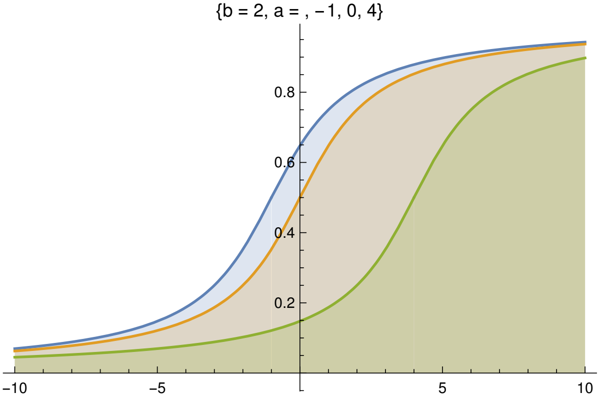
Distribución Cauchy

Función de densidad



Distribución Cauchy

Función de distribución acumulada



Distribución Pareto

$$X \sim \text{Pareto}(\alpha, \beta)(\alpha > 0, \beta > 0$$

$$f(x) = \begin{cases} \frac{\alpha\beta^{\alpha}}{x^{(\alpha+1)}} & \text{si } x > \beta \\ 0 & \text{si} \end{cases}$$

$$E(X) = \frac{\alpha\beta}{\alpha - 1} \text{ para } \alpha > 1$$

$$Var(X) = \frac{\alpha\beta^2}{(\alpha - 1)^2(\alpha - 2)} \text{ para } \alpha > 1$$

Distribución Weibull

Sea $X_{1:n}$ una muestra aleatoria de una distribución Weibull con parámetro de escala $\lambda = 1$ y parámetro de forma α, $X \sim \text{Weibull}(\alpha, 1)$

$$f(x_{i:1}, \alpha) = \alpha x_i^{\alpha-1} e^{-x_i^{\alpha}}$$