Otras Distribuciones Multivariadas

José A. Perusquía Cortés

Análisis Multivariado Semestre 2023-2

Distribuciones Elípticas y Esféricas

- **Definición.**

  Decimos que $\mathbf{x}$ tiene una distribución elíptica denotado por $\mathbf{x} \sim EC(\nu, \Lambda, g)$ si su densidad está dada por

$$|\Lambda|^{-\frac{1}{2}} g\left[(\mathbf{x}-\nu)^T \Lambda^{-1} (\mathbf{x}-\nu)\right]$$

- Donde
  - $\Lambda > 0$
  - $g(\cdot) \geq 0$ y $\int g(\mathbf{y}^T\mathbf{y})d\mathbf{y} = 1$

- **Observación 1**

  Si $\mathbf{C}$ es una matriz no singular tal que $\mathbf{C}^T\Lambda^{-1}\mathbf{C} = \mathbf{I}$ y usamos la transformación $\mathbf{x} - \nu = \mathbf{C}\mathbf{y}$ entonces $\mathbf{y}$ tiene densidad

$$g(\mathbf{y}^T\mathbf{y})$$

- **Observación 1**

  Si $\mathbf{C}$ es una matriz no singular tal que $\mathbf{C}^T\Lambda^{-1}\mathbf{C} = \mathbf{I}$ y usamos la transformación $\mathbf{x} - \nu = \mathbf{C}\mathbf{y}$ entonces $\mathbf{y}$ tiene densidad

$$g(\mathbf{y}^T\mathbf{y})$$

- Donde las curvas de nivel son esferas centradas en el origen. La clase de estas densidades es conocida como distribuciones esféricas y denotamos $\mathbf{y} \sim S(g)$

- **Observación 2**

  Una densidad esférica puede representarse en coordenadas polares mediante la transformación

$$
\begin{aligned}
y_1 &= r\sin(\theta_1)\\
y_2 &= r\cos(\theta_1)\sin(\theta_2)\\
y_3 &= r\cos(\theta_1)\cos(\theta_2)\cos(\theta_3)\\
&\vdots\\
y_{p-1} &= r\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{p-2})\sin(\theta_{p-1})\\
y_p &= r\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{p-2})\cos(\theta_{p-1})
\end{aligned}
$$

- Donde
  - Para $i \in 1,\ldots,p-2$ se tiene que $-\frac{\pi}{2} < \theta_i \leq \frac{\pi}{2}$ y $-\pi < \theta_{p-1} \leq \pi$
  - $0 \leq r < \infty$

- Así la densidad de $(R, \Theta_1, \ldots, \Theta_{p-1})$ es (Tarea)

$$r^{p-1}\cos(\theta_1)^{p-2}\cos(\theta_2)^{p-3}\cdots\cos(\theta_{p-2})g(r^2)$$

- Así la densidad de $(R, \Theta_1, \ldots, \Theta_{p-1})$ es (Tarea)

$$r^{p-1}\cos(\theta_1)^{p-2}\cos(\theta_2)^{p-3}\cdots\cos(\theta_{p-2})g(r^2)$$

- **Observación 3**

  Las variables aleatorias $R, \Theta_1, \ldots, \Theta_{p-1}$ son independientes

- Así la densidad de $(R, \Theta_1, \ldots, \Theta_{p-1})$ es (Tarea)

$$r^{p-1}\cos(\theta_1)^{p-2}\cos(\theta_2)^{p-3}\cdots\cos(\theta_{p-2})g(r^2)$$

- **Observación 3**

  Las variables aleatorias $R, \Theta_1, \ldots, \Theta_{p-1}$ son independientes

- **Observación 4** (Tarea)

  La distribución marginal de R es

$$\frac{2\pi^{\frac{p}{2}}r^{p-1}g(r^2)}{\Gamma\left(\frac{p}{2}\right)}$$

- **Observación 5** (Tarea)

  La distribución marginal de Θ_i para $i \in 1,\ldots,p-2$

$$\frac{\Gamma\left(\frac{p-i}{2}\right)\cos(\theta_i)^{p-i-1}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-i-1}{2}\right)}$$

- **Observación 5** (Tarea)

  La distribución marginal de Θ_i para $i \in 1, \ldots, p-2$

$$\frac{\Gamma\left(\frac{p-i}{2}\right)\cos(\theta_i)^{p-i-1}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-i-1}{2}\right)}$$

  La distribución marginal de Θ_{p-1} es

$$\frac{1}{2\pi}$$

- **Definición**

  Un vector aleatorio $\mathbf{y}$ tiene una distribución esférica si para toda $\mathbf{O} \in \mathbb{O}(n)$ se tiene que $\mathbf{y} \overset{d}{=} \mathbf{O}\mathbf{y}$

- **Definición**

  Un vector aleatorio $\mathbf{y}$ tiene una distribución esférica si para toda $\mathbf{O} \in \mathbb{O}(n)$ se tiene que $\mathbf{y} \overset{d}{=} \mathbf{O}\mathbf{y}$

- **Teorema 1**

  Un vector aleatorio $\mathbf{y}$ tiene una distribución esférica si y solo si su función característica $\phi_{\mathbf{y}}(t)$ satisface alguna de estas condiciones equivalentes

  1. $\phi_{\mathbf{y}}(\mathbf{O}t^t) = \phi_{\mathbf{y}}(t)$ para toda $\mathbf{O} \in \mathbb{O}(n)$
  2. Existe una función escalar (generador característico) $\varphi(\,\cdot\,)$ tal que $\phi_{\mathbf{y}}(t) = \varphi(t^T t)$

- **Ejemplos**

  - Familia de Kotz.

$$g(u) = C_n u^{N-1} \exp\left(-ru^s\right)$$

- Donde

  - $r, s > 0$
  - $2N + n > 2$
  - C_n es la constante de normalización

- **Observación**

  La densidad está dada por

$$C_n \,|\, \Lambda \,|^{-\frac{1}{2}} \left[(\mathbf{x}-\nu)^T \Lambda^{-1} (\mathbf{x}-\nu)\right]^{N-1} \exp\left\{-r\left[(x-\nu)^T \Lambda^{-1} (x-\nu)\right]^s\right\}$$

- **Observación**

  La densidad está dada por

$$C_n |\Lambda|^{-\frac{1}{2}} \left[(\mathbf{x}-\nu)^T\Lambda^{-1}(\mathbf{x}-\nu)\right]^{N-1} \exp\left\{-r\left[(x-\nu)^T\Lambda^{-1}(x-\nu)\right]^s\right\}$$

- Si $N = 1$, $s = 1$ y $r = \dfrac{1}{2}$ obtenemos la distribución normal multivariada.
- Familia útil cuando el supuesto de normalidad no es aplicable

- Familia de Pearson del tipo VII (incluye a la distribución t y a la Cauchy)

$$g(u) = C_n \left(1 + \frac{u}{m}\right)^{-N}$$

- Donde
  - $m > 0$
  - $N > \frac{n}{2}$
  - C_n es la constante de normalización

Distribución t multivariada

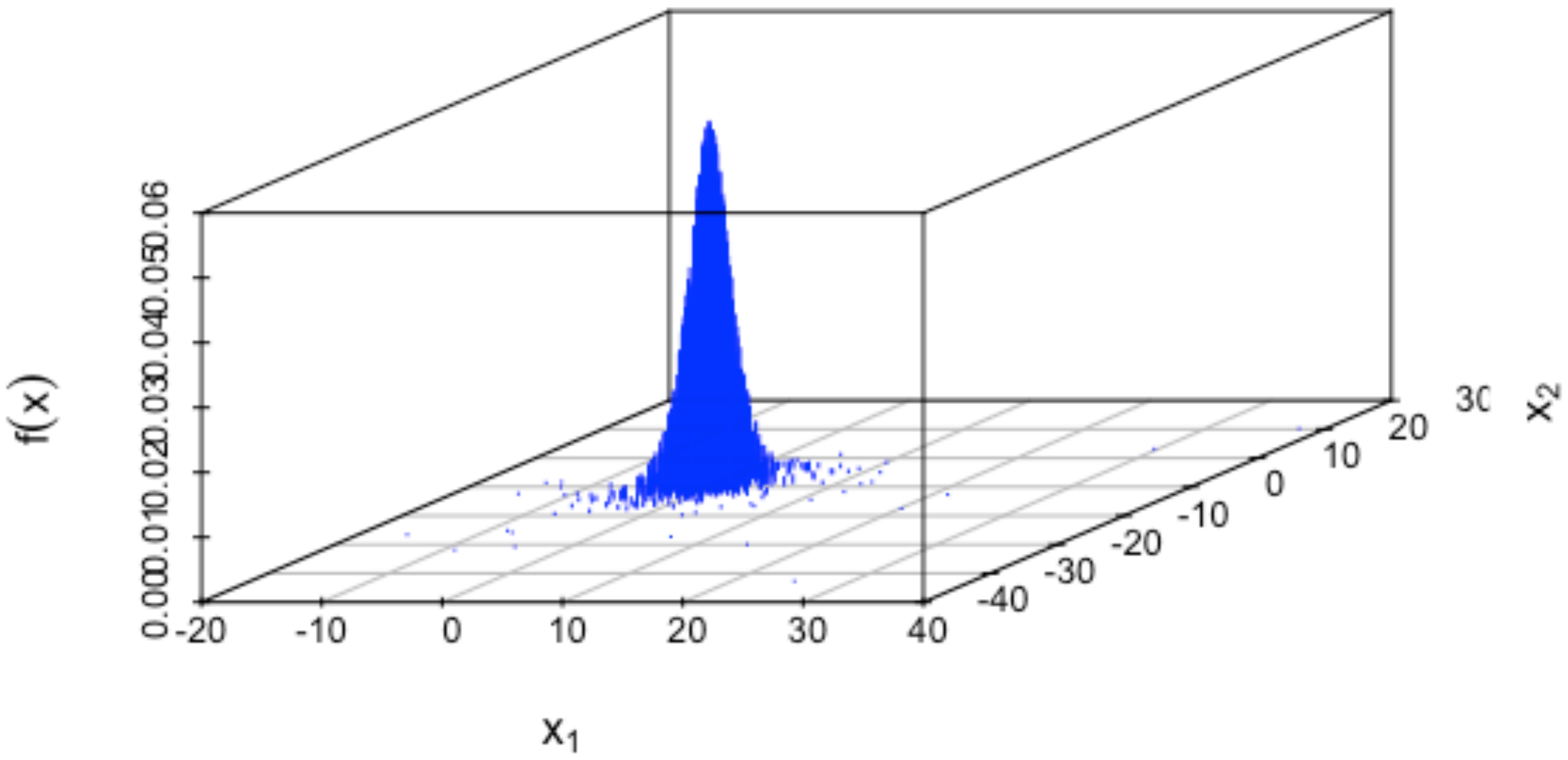
- Caso particular de la familia de Pearson del tipo VII con $N = \frac{1}{2}(n+m)$
- Denotado por $\mathbf{x} \sim Mt_n(m, \nu, \Lambda)$ y con densidad

$$\frac{\Gamma\left(\frac{n+m}{2}\right)}{(\pi m)^{\frac{n}{2}}\Gamma\left(\frac{m}{2}\right)} |\Lambda|^{-\frac{1}{2}} \left(1 + m^{-1}(\mathbf{x}-\nu)^T\Lambda^{-1}(\mathbf{x}-\mu)\right)^{-\frac{n+m}{2}}$$

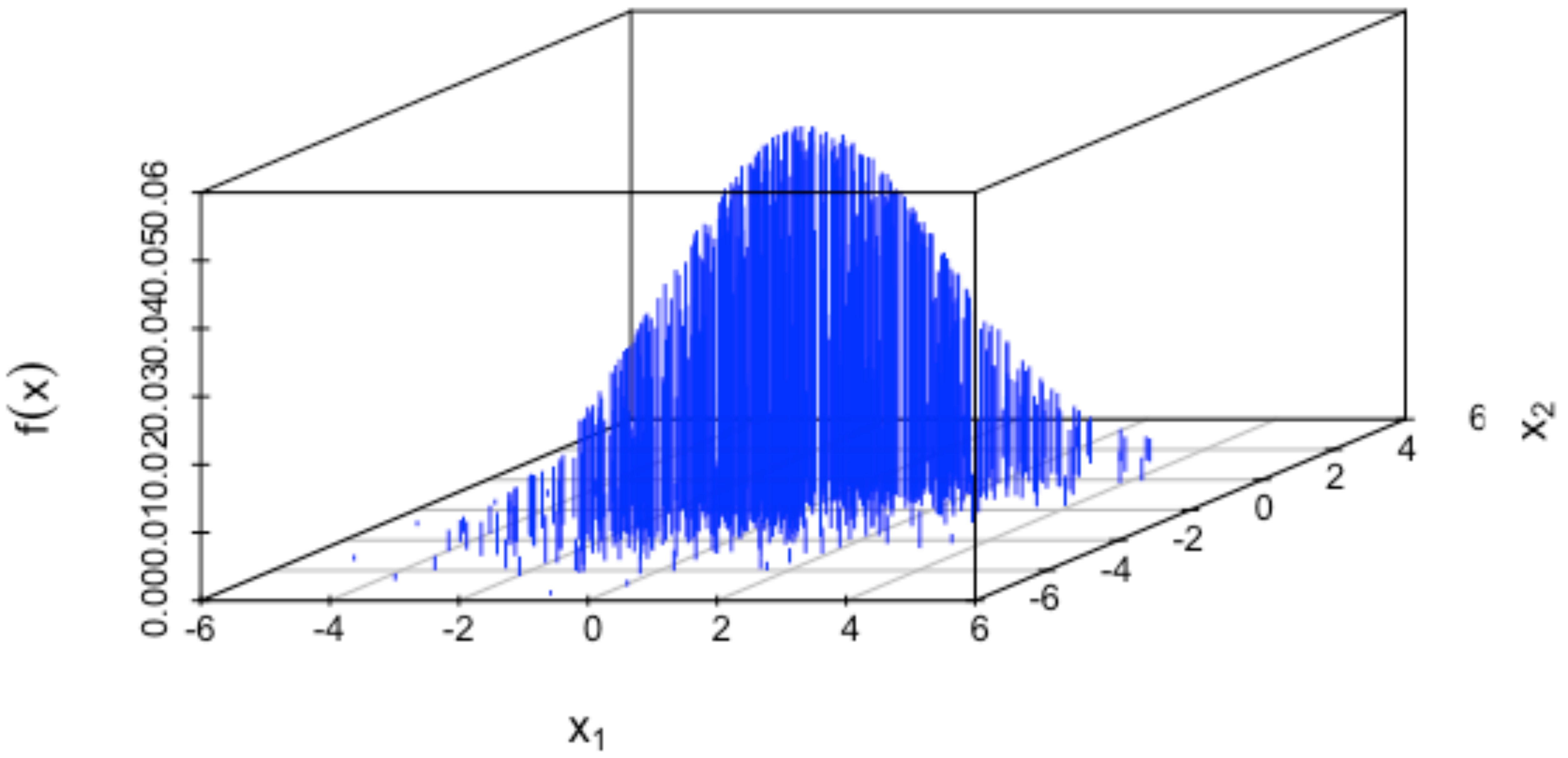
- Para la distribución t multivariada en **R** podemos usar la librería **mvtnorm** (por default en escala logarítmica)

  - **dmvt()** : Evaluar la densidad
  - **pmvt()** : Evaluar la distribución.
  - **qmvt()** : Obtener los cuantiles.
  - **rmvt()** : Obtener una muestra.

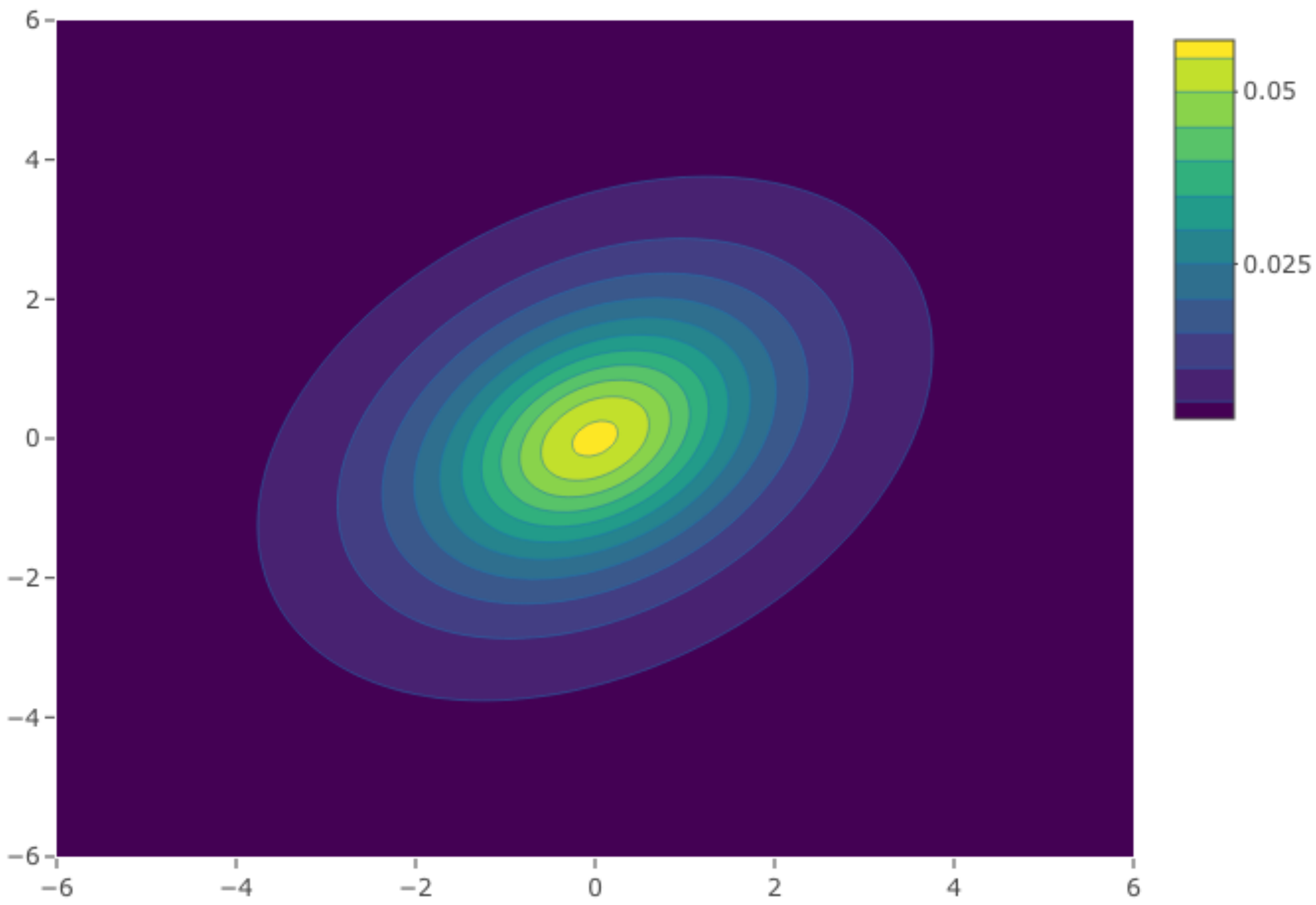
Distribución t (df=2)



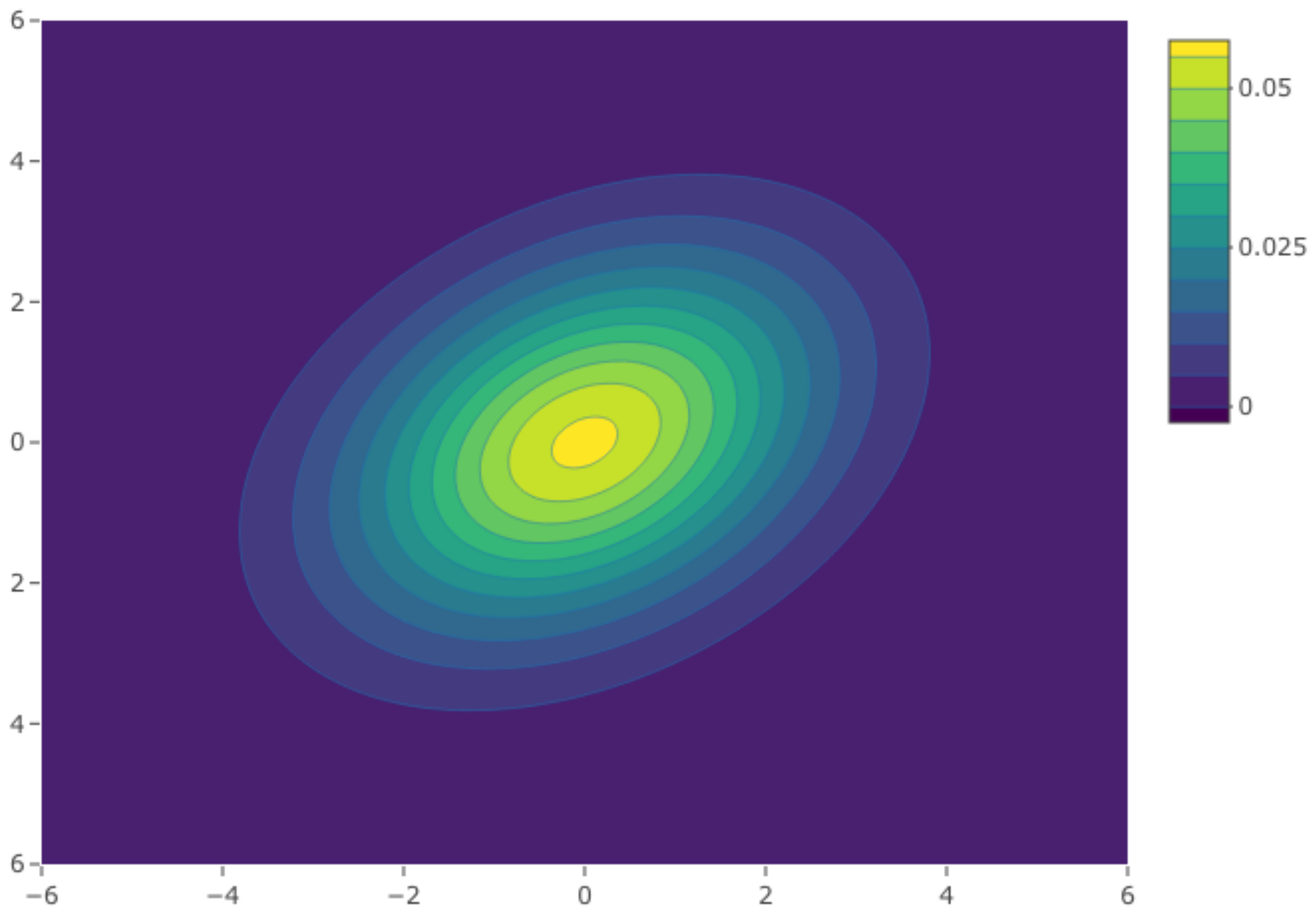
Distribución normal



Distribución t (df=2)



Distribución normal



Distribución Cauchy multivariada

- Caso particular de la distribución t multivariada con $m = 1$

Distribución Cauchy

- Caso particular de la distribución t multivariada con $m = 1$
- Denotado por $\mathbf{x} \sim MC_n(\nu, \Lambda)$ y con densidad

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} |\Lambda|^{-\frac{1}{2}} \left(1 + (\mathbf{x} - \nu)^T \Lambda^{-1} (\mathbf{x} - \mu)\right)^{-\frac{n+1}{2}}$$

Distribución Cauchy

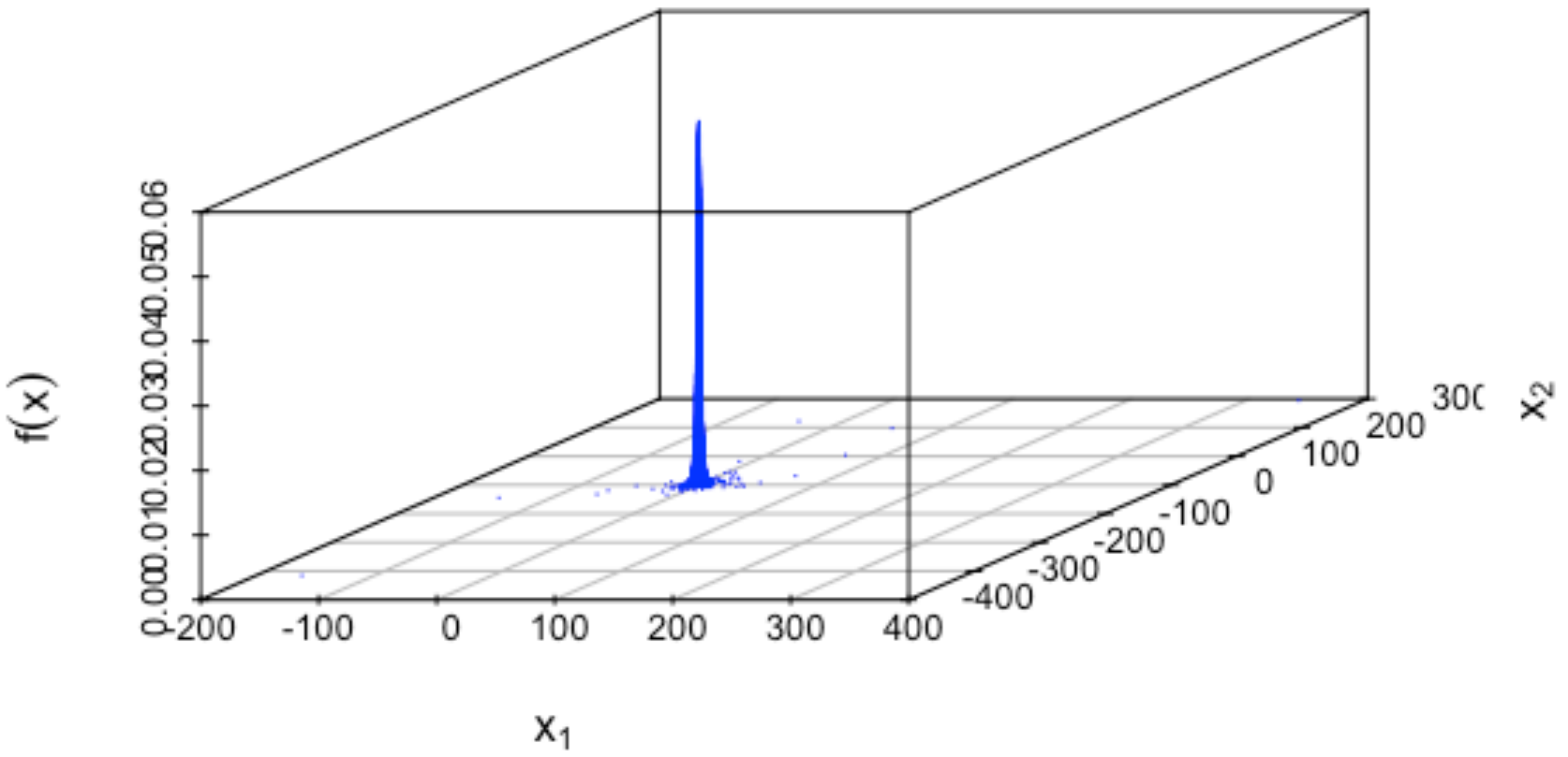
- Caso particular de la distribución t multivariada con $m = 1$
- Denotado por $\mathbf{x} \sim MC_n(\nu, \Lambda)$ y con densidad

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} |\Lambda|^{-\frac{1}{2}} \left(1 + (\mathbf{x} - \nu)^T \Lambda^{-1} (\mathbf{x} - \mu)\right)^{-\frac{n+1}{2}}$$

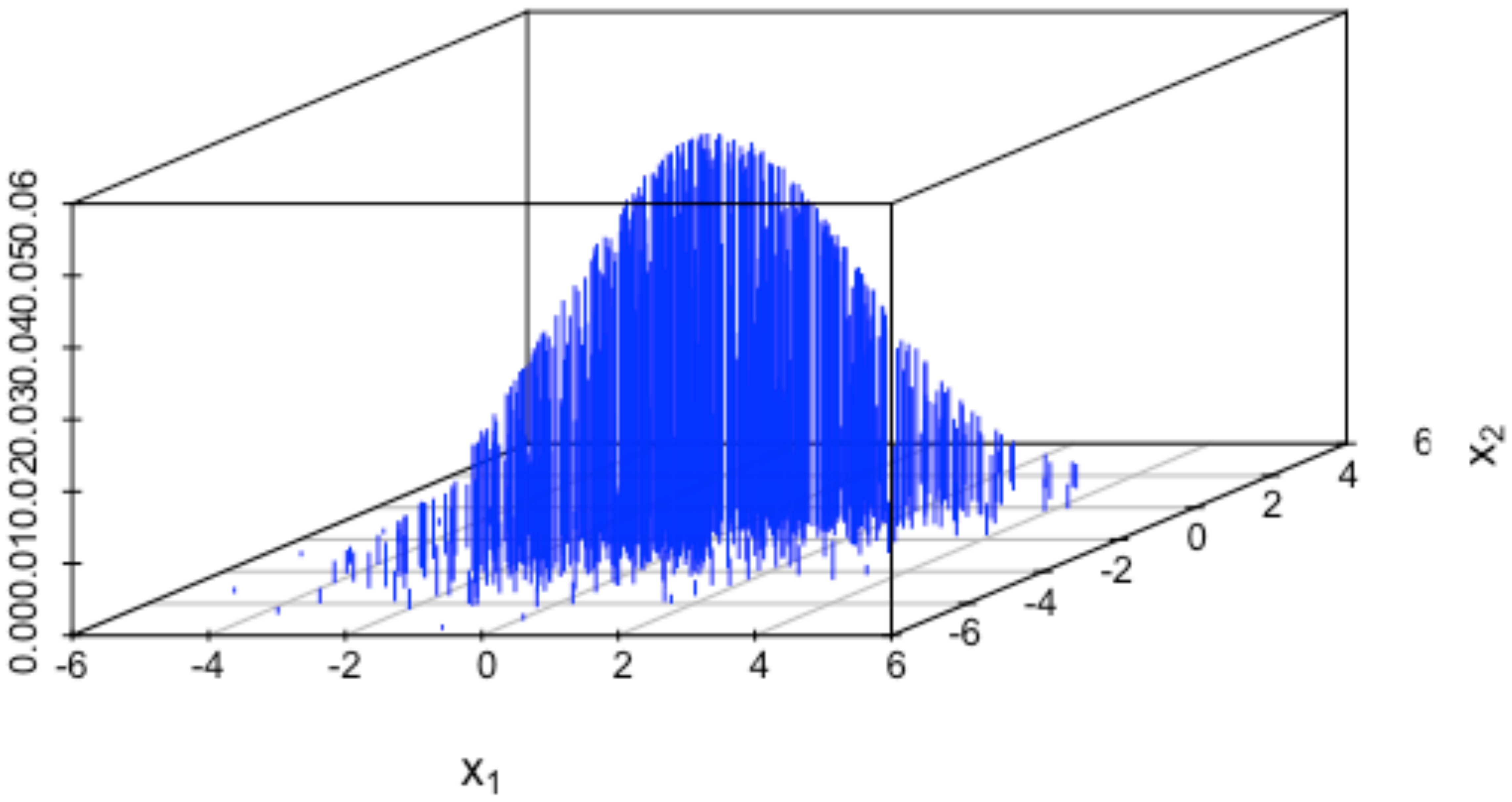
- El caso particular $\mathbf{x} \sim MC_n(\mathbf{0}, \mathbf{I})$

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \left(1 + \mathbf{x}^T \mathbf{x}\right)^{-\frac{n+1}{2}}$$

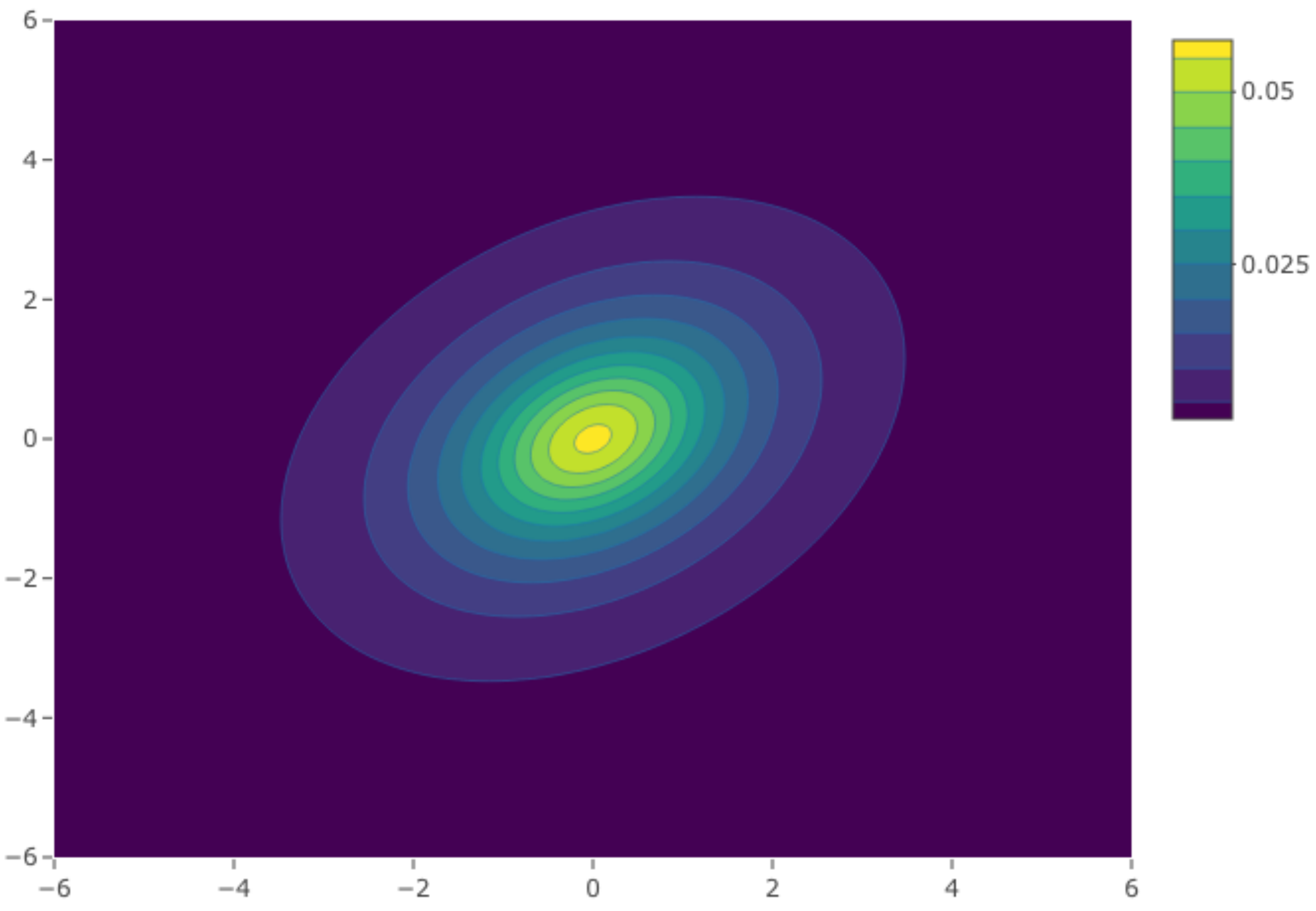
Distribución Cauchy



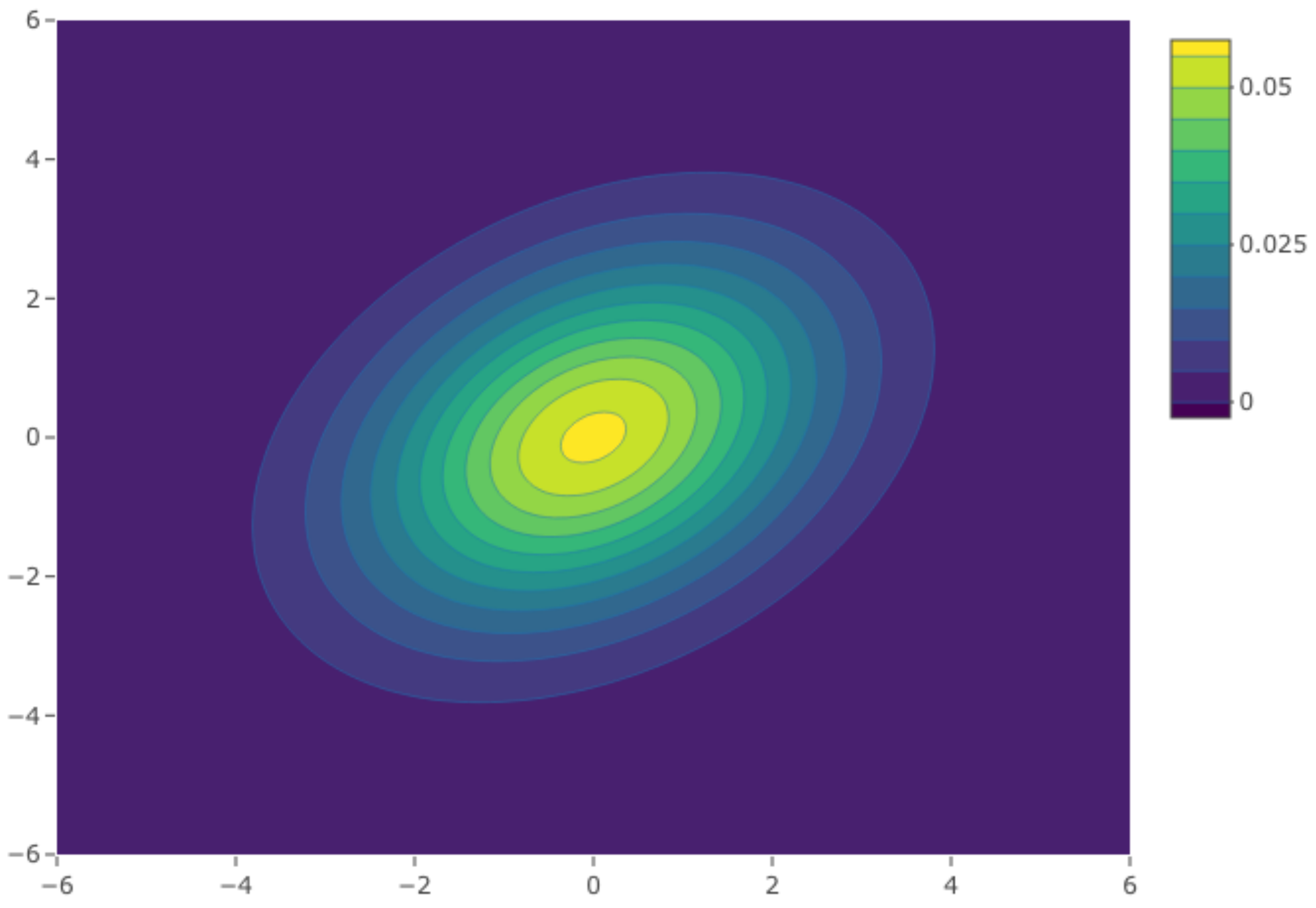
Distribución normal



Distribución Cauchy



Distribución normal



Distribución Dirichlet

Distribución Dirichlet

- Distribución continua y generalización de la distribución beta

Distribución Dirichlet

- Distribución continua y generalización de la distribución beta
- Soporte $\left\{ \mathbf{x} \in [0,1]^p : \sum_i x_i = 1 \right\}$ (simplex (p-1)-dimensional)

Distribución Dirichlet

- Distribución continua y generalización de la distribución beta
- Soporte $\left\{ \mathbf{x} \in [0,1]^p : \sum_i x_i = 1 \right\}$ (simplex (p-1)-dimensional)

$$f(\mathbf{x}) = \frac{\Gamma\left(\sum_{i=1}^{p} \alpha_i\right)}{\prod_{i=1}^{p} \Gamma(\alpha_i)} \prod_{i=1}^{p} x_i^{\alpha_i - 1}$$

- Donde
  - $\alpha_i > 0 \ \forall i$ (parámetros de concentración)

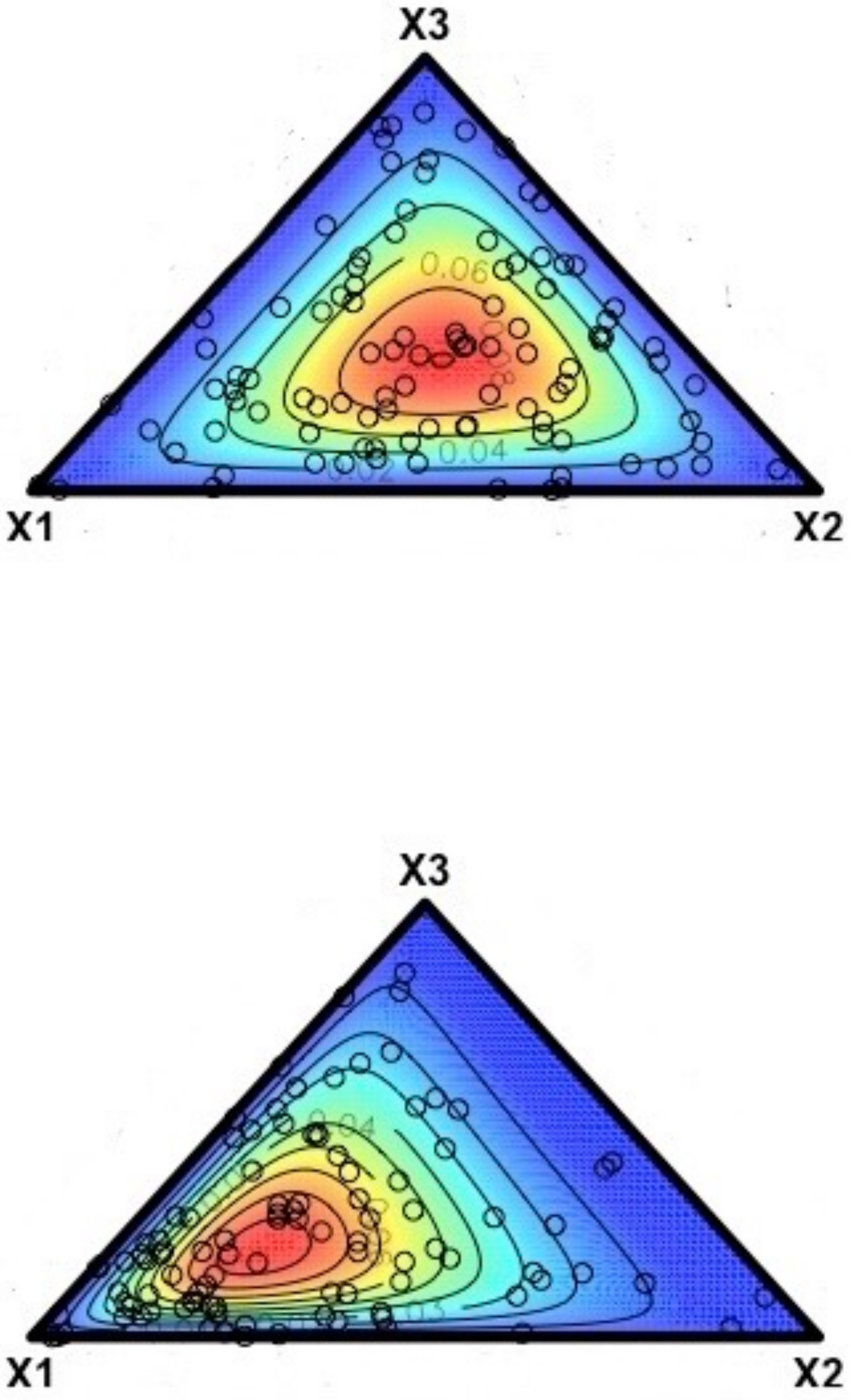
- **Proposición** **(Tarea)**

Sean $y_i \sim Ga(\alpha_i, \theta)$ una colección de variables aleatorias independientes entonces $V = \sum_i y_i \sim Ga(\alpha_0, \theta)$ y

$$(x_1, \ldots, x_p) = \left(\frac{y_1}{V}, \ldots, \frac{y_p}{V}\right) \sim Dir(\alpha_1, \ldots, \alpha_p)$$

- Donde

  - $\alpha_0 = \sum_i \alpha_i$


X3
X1
X2
0.06
0.04
0.02
X3
X1
X2
0.0018
X3
X1
X2
0.04
X3
X1
X2

Distribución Multinomial

Distribución Multinomial

- Distribución discreta y generalización de la distribución binomial

- Distribución discreta y generalización de la distribución binomial
- Soporte $\left\{ \mathbf{x} \in \mathbb{N}^p : \sum_i x_i = N \right\}$ y con función masa

Distribución Multinomial

- Distribución discreta y generalización de la distribución binomial
- Soporte $\left\{ \mathbf{x} \in \mathbb{N}^p : \sum_i x_i = N \right\}$ y con función masa

$$\mathbb{P}(\mathbf{x} = \mathbf{n}) = \frac{N!}{n_1! \cdots n_p!} \prod_{i=1}^{p} a_i^{n_i}$$

- Donde
  - $N \in \mathbb{N}$
  - $a_i \in (0,1)\ \forall i$ y $\sum_{i=1}^{p} a_i = 1$

Otras

Otras

- Distribución Bernoulli multivariada
- Distribución Poisson multivariada
- Distribución Pareto multivariada
- Distribución exponencial multivariada
- Etc.