REGRESIÓN BAYESIANA

Equipo 3

Ana Saraí Dávila Martínez 1986209

Jesús Mauricio Rodríguez Pacheco 1583211

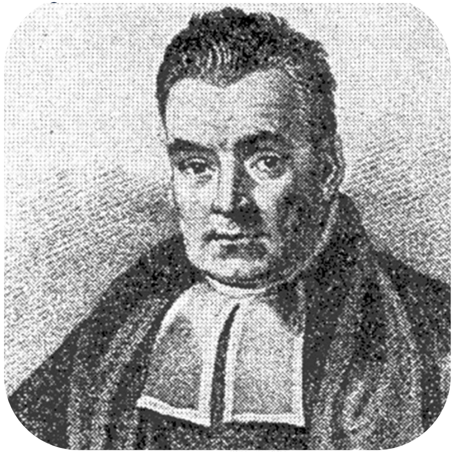
Milton Aldair Martínez Acosta 1870354

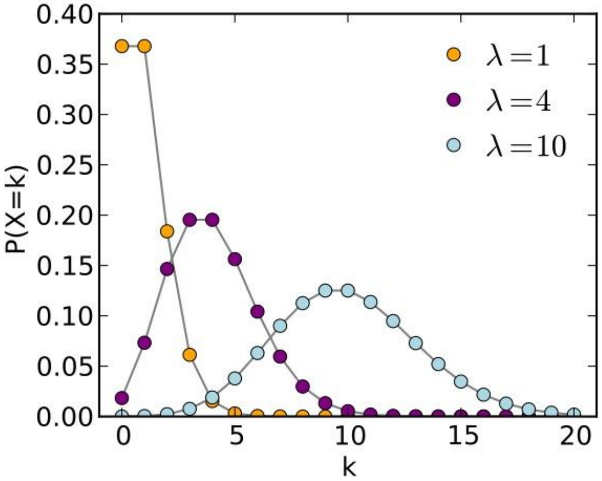
ESTADÍSTICA BAYESIANA

La inferencia bayesiana está basada en la distribución de probabilidad del parámetro dado los datos (distribución a posteriori de probabilidad P(Θ|y), en lugar de la distribución de los datos dado el parámetro.

La aplicación del modelo bayesiano en modelos de regresión, ya sea estándar, logística, Poisson, etc, sigue el esquema general de la estadística bayesiana:

- Definir la distribución a priori correspondiente para los parámetros.
- Determinar la verosimilitud de los datos.
- Aplicar el teorema de Bayes para actualizar la distribución a priori en forma de distribución a posteriori.

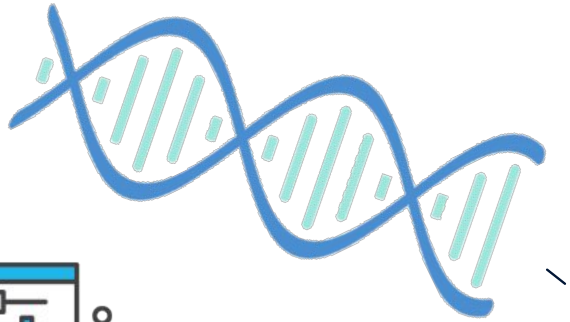




APLICACIONES

Tiene aplicaciones a distintas áreas del conocimiento como:

- Medicina
- Biología
- Astronomía
- Economía
- Planificación de proyectos sociales



TEOREMA DE BAYES

La probabilidad de *B* condicionada a *A*, se expresa como:

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

El evento B no es observable, A es el evento observable y pasa a ser el espacio muestral reducido de este experimento.

- Cambiando el rol de los evento, se tiene que:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(B) * P(A/B)$$

- Por la regla de multiplicación

$$P(A) = \sum_{i=1}^{n} P(B_i) * P(A/B_i)$$

- Y por la expresión:

$$P(B_i/A) = \frac{P(A \cap B_i)}{P(A)}$$

$$\boxed{P(B_i/A) = \frac{P(B_i) * P(A/B_i)}{\sum_{i=1}^{n} P(B_i) * P(A/B_i)}}$$

Representa la formulación del teorema de Bayes, que fue publicado en 1763

PROCESO DE INFERENCIA

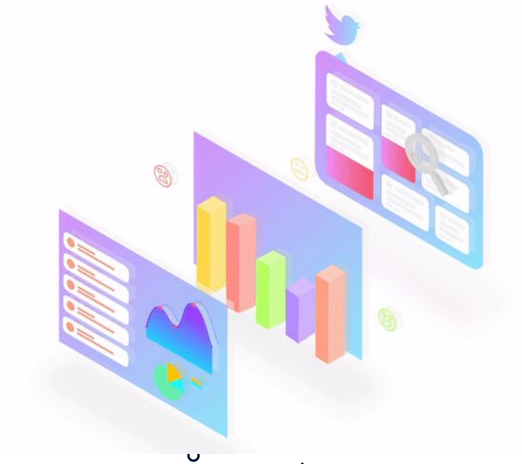
La información que contienen los datos en relación a los parámetros es expresada en la función de verosimilitud, y que se representa como p(y/θ). Esta información es utilizada para actualizar la distribución a priori p(θ).

El teorema de Bayes establece como actualizar la distribución a posteriori de p(θ). En este sentido, la distribución a posteriori es proporcional a la distribución a priori multiplicada por la función de verosimilitud, es decir:

$$P(\theta/y) = \frac{P(\theta) * P(y/\theta)}{\int P(\theta) * P(y/\theta) d\theta}$$

o alternativamente:

$$posterior \propto prior \times verosimilitud$$



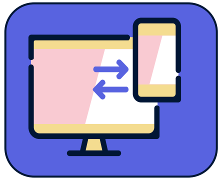
PROCESO DE INFERENCIA

MCMC

Markov Chain Monte Carlo: son métodos de integración que utilizan Cadenas de Markov

METROPOLIS-HASTINGS ALGORITMO

Para construir estas cadenas se utiliza el procedimiento desarrollado por Metropolis et al. (1953) y Hastings (1970)



DISTRIBUCIÓN A POSTERIORI

Se pueden analizar cualquier característica (media, momentos, cuantiles, percentiles, etc).

PROCESO DE INFERENCIA

Considerando la esperanza:

$$E(f(\theta/y)) = \frac{\int f(\theta)P(\theta) * P(y/\theta)d(\theta)}{\int P(\theta) * P(y/\theta)d\theta}$$

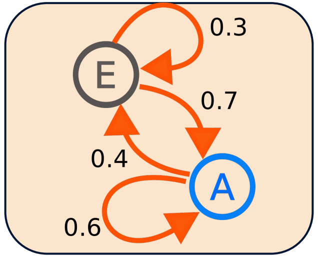
Generalizando, dado un vector X con distribución de probabilidad π(·), se trata de evaluar el valor esperado para una función f(·) de interés, es decir:

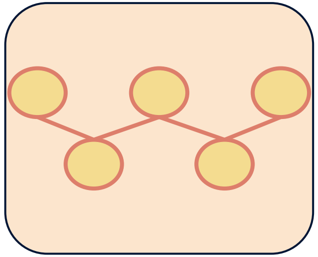
$$E(f(X)) = \frac{\int f(x)\pi(x)dx}{\int \pi(x)dx}$$

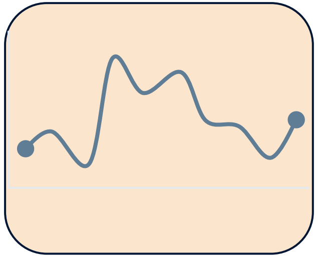
Con la integración de Monte Carlo, este valor se estima a través del promedio de las muestras realizadas:

$$E(f(X)) = \frac{1}{n}\sum_{t=1}^{n} f(x_t)$$

CADENAS DE MARKOV







¿QUÉ ES?

Una Cadena de Markov es una secuencia de variables aleatorias

$$X_0, X_1, X_2, \ldots$$

PROPIEDAD DE MARKOV

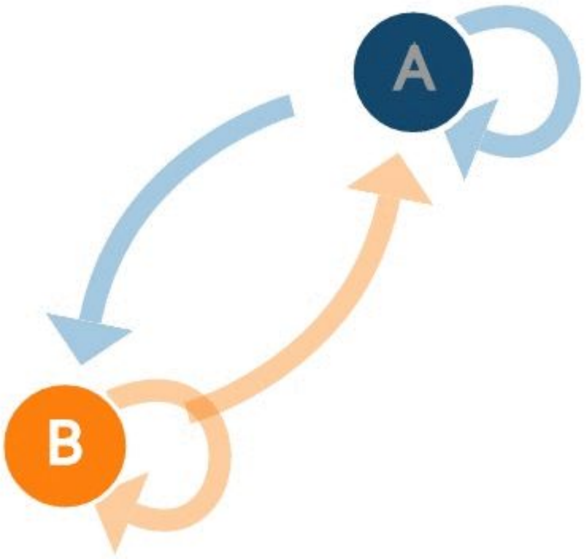
Sólo se tiene en cuenta el estado actual de la cadena

KERNEL DE TRANSICIÓNI

La distribución $P(\cdot / \cdot)$ se supone independiente de t,

CADENAS DE MARKOV

MEDIA ERGÓDICA



- Al aumentar t, los valores de la muestra Xt irán asemejándose cada vez más a los valores muestrales de ϕ(·). Así, los valores $X_{m+1}, X_{m+2}, \ldots, X_n$, se aproximarán a los de la distribución ϕ(·). Podemos estimar, pues, la E(f(X)) como:

$$\boxed{E(f(X)) = \bar{f} = \frac{1}{n-m}\sum_{t=m+1}^{n} f(x_t)}$$

Esta expresión muestra como una Cadena de Markov puede ser utilizada para estimar E(f(X))

REGRESIÓN LINEAL BAYESIANA

- En un modelo de regresión simple, es decir $y_i = \alpha + \beta x_i + \epsilon_i$, se necesita definir la distribución conjunta a priori sobre $\boldsymbol{\alpha}$,$\boldsymbol{\beta}$ y $\boldsymbol{\sigma}$, es decir el intercepto, pendiente de la recta y el error.
- Si suponemos que

$$\alpha \sim N(-\infty, \infty) \qquad \beta \sim N(-\infty, \infty) \qquad log(\sigma) \sim N(-\infty, \infty)$$

Se pueden aproximar los resultados de la regresión clásica. Se necesita considerar el logaritmo de la desviación estándar, dado que la varianza debe ser positiva. La distribución a priori es equivalente a una densidad que es proporcional a $\frac{1}{\sigma^2}$.

REGRESIÓN LINEAL BAYESIANA

Dado que los residuos se distribuyen como una normal, $\epsilon_i \sim N(0, \sigma)$, la esperanza de la regresión será:

$$\boxed{E(y) = \alpha + \beta x}$$

Y, por tanto, se puede deducir que $y \sim N(\alpha + \beta x, \sigma)$.

La función de verosimilitud será:

$$\boxed{f(y_i) = \frac{1}{\sqrt{2\pi\sigma}} exp[\frac{1}{\sigma^2}(y_i - (\alpha + \beta x_i))^2]}$$

Y la conjunta, $\prod_{i=1}^{n} f(y_i)$,será por tanto una distribución normal multivariante de dimensión *n*.

Las distribuciones a posteriori de $\boldsymbol{\alpha}$ y de $\boldsymbol{\beta}$ son t-Student con n-2 grados de libertad, y $\boldsymbol{\sigma}^2$ se distribuye como chi cuadrado inversa.

EJEMPLO

Librerías utilizadas

PyMC3

Definición de modelos

ArviZ

Análisis de modelos bayesianos

NumPy

Trabaja con matrices multidimensionales

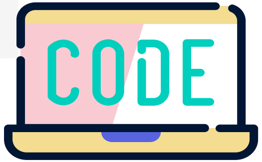
matplotlib

Generación de gráficos

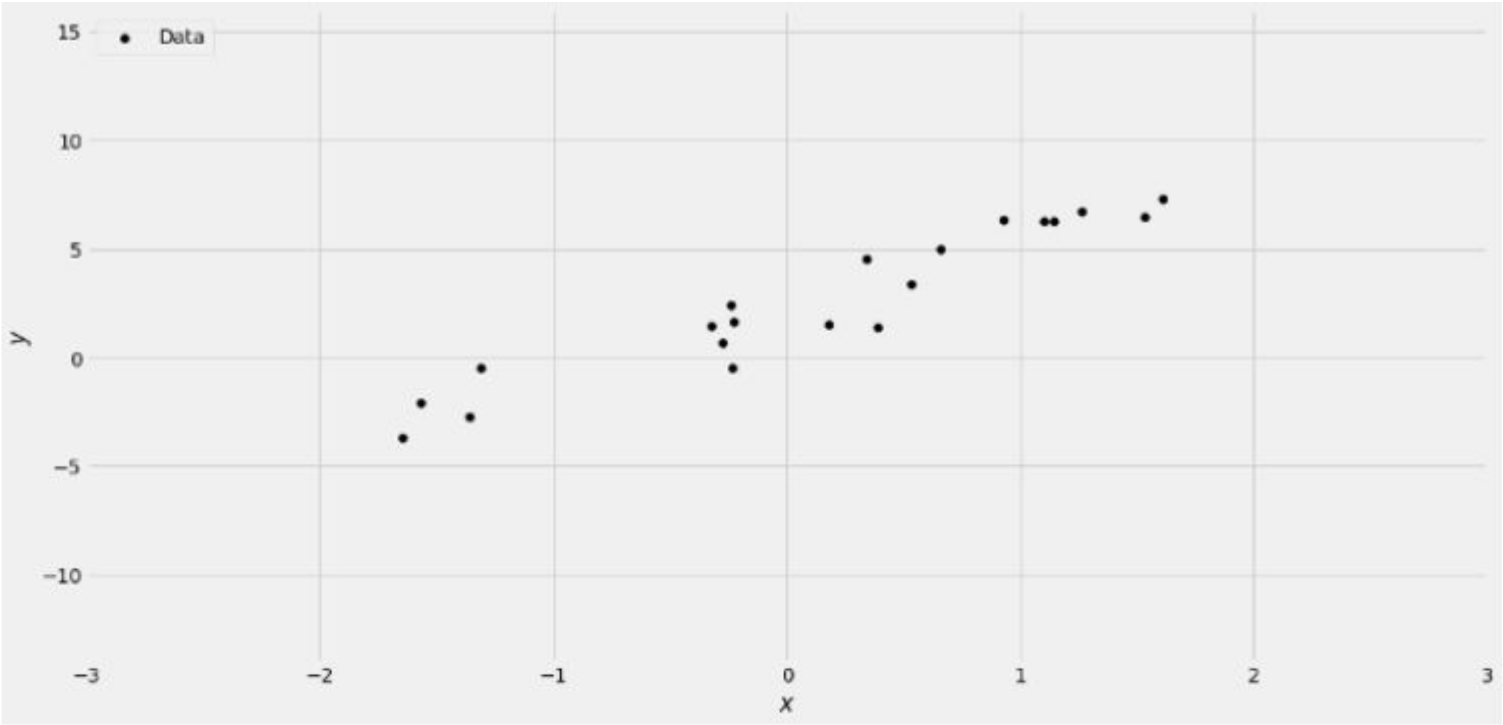
https://github.com/AnaDavila1/Mineria-de-datos-FCFM/blob/main/Regresi%C3%B3n%20Bayesiana%20Ejemplo.ipynb

CONJUNTO DE DATOS INICIALES

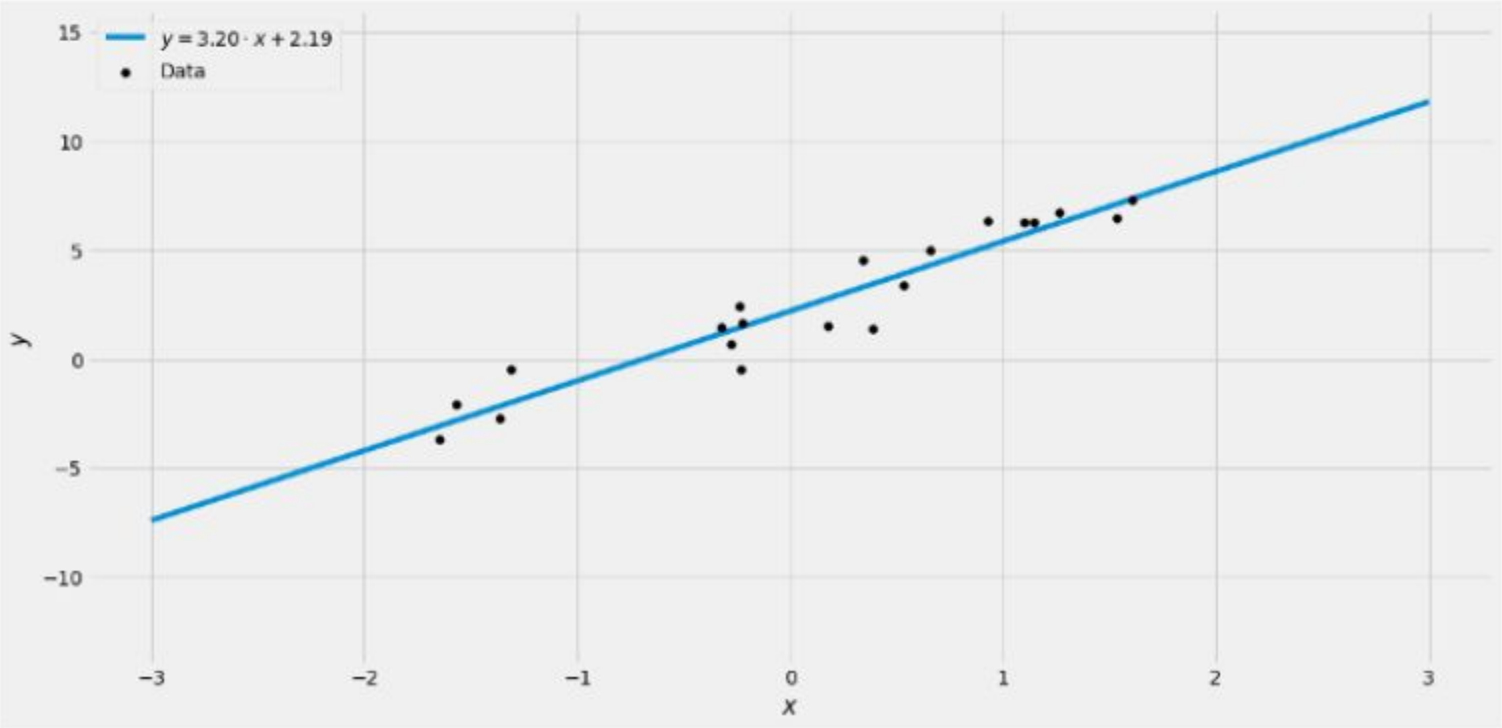
```
x = [
   -1.64934805,  0.52925273,  1.10100092,  0.38566793, -1.56768245,
    1.26195686,  0.92613986, -0.23942803,  0.33933045,  1.14390657,
    0.65466195, -1.36229805, -0.32393554, -0.23258941,  0.17688024,
    1.60774334, -0.22801156,  1.53008133, -1.31431042, -0.27699609
] # inputs
y = [
   -3.67385666,  3.37543275,  6.25390538,  1.41569973, -2.08413872,
    6.71560158,  6.32344159,  2.40651236,  4.54217349,  6.25778739,
    4.98933806, -2.69713137,  1.45705571, -0.49772953,  1.50502898,
    7.27228263,  1.6267433 ,  6.43580518, -0.50291509,  0.65674682
] # outputs
```



VISUALIZACIÓN DE LOS PARES ORDENADOS



RESULTADO DE UNA REGRESIÓN LINEAL



SUPOSICIONES

$$a, b \sim \mathcal{N}(0, 16)$$

a -> pendiente de la recta

b-> intersección con el eje y

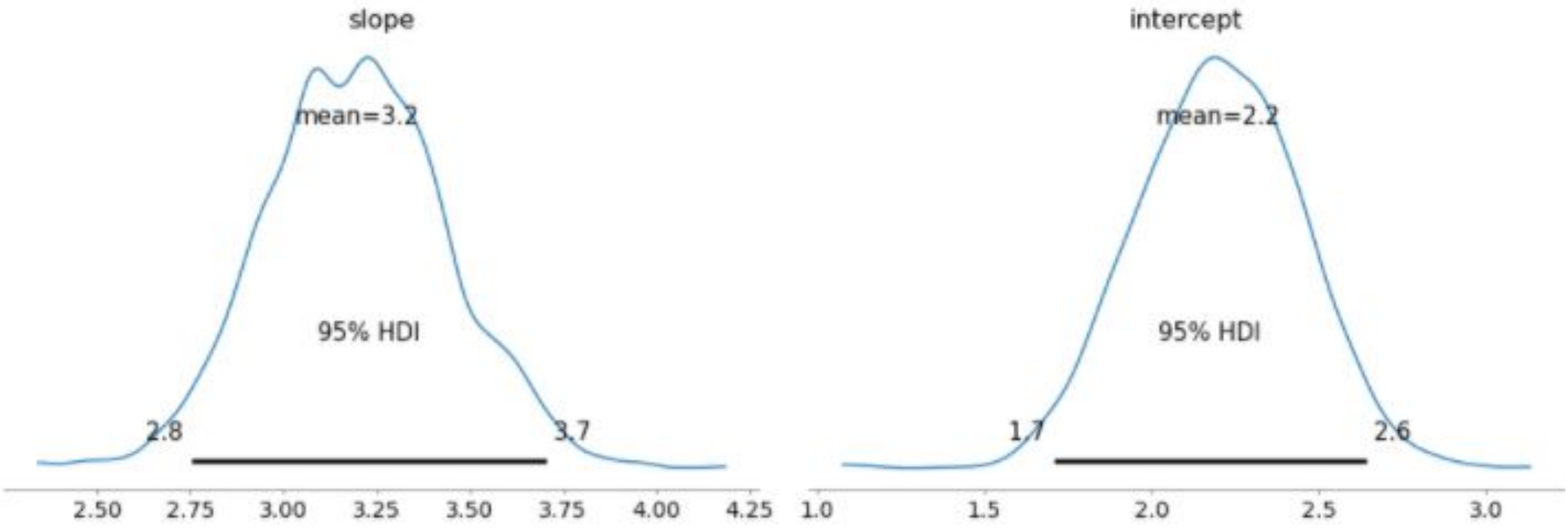
$$\sigma^2 \sim \mathrm{Exp}(1)$$

σ^2-> Error

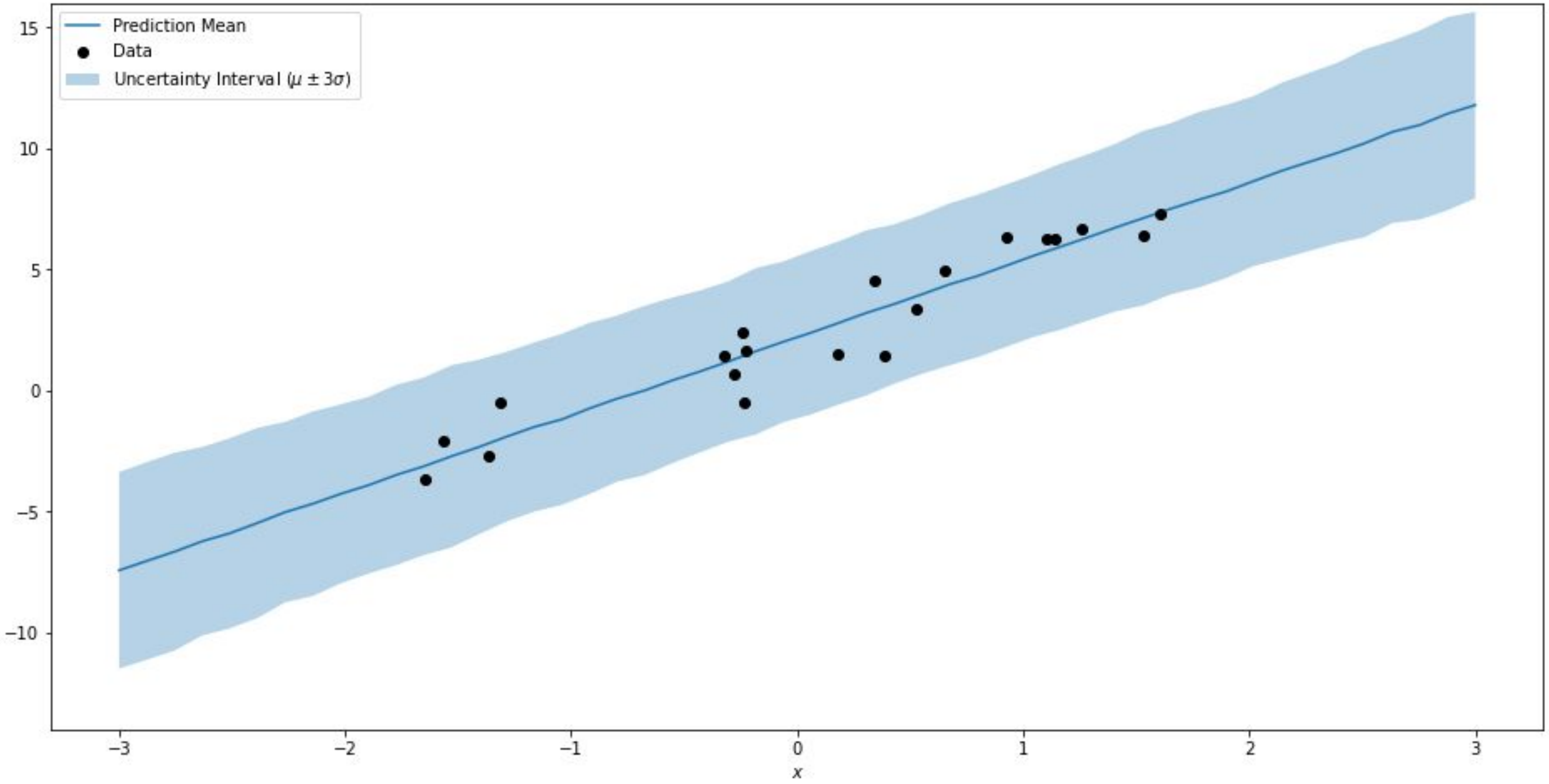
MODELO EN PYTHON

```
with pm.Model() as predictive_model:
    a = pm.Normal('slope', 0, 16)
    b = pm.Normal('intercept', 0, 16)
    s = pm.Exponential('error', 1)

    x_ = pm.Data('features', x) # almacenador de datos

    obs = pm.Normal('observation', a*x_ + b, s, observed=y)

    trace = pm.sample()
```

DISTRIBUCIONES DE LOS PARÁMETROS



RESULTADO DE UNA REGRESIÓN BAYESIANA



REFERENCIAS

- Parra, F. (2019). Regresión bayesiana. marzo 09, 2021, de bookdown Sitio web: https://bookdown.org/content/2274/modelos-con-variables-cualitativas.html#regresion-bayesiana
- Paredes, O.. (2013). Inferencia bayesiana. En Regresión Lineal por Medio del Análisis Bayesiano(pp. 45). México.
- Regresión lineal bayesiana en Python a través de PyMC3. ichi.pro. https://ichi.pro/es/regresion-lineal-bayesiana-en-python-a-traves-de-pymc3-188618526786065

MUCHAS GRACIAS