Formative 3 – Probability Distributions, Bayesian Probability, and Gradient Descent Implementation

$$y = mx + b$$

Using data points (1,3) and (3,6), with:

- Initial $m = -1$
  Initial $b = 1$
  → Learning rate $\alpha = 0.1$

Cost function is Mean Squared Error (MSE)

Predicted Value:

$$\hat{y}_i = mx_i + b$$

Gradients (for MSE):

$$\frac{\partial J}{\partial m} = -\frac{2}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)\,x_i$$

$$\frac{\partial J}{\partial b} = -\frac{2}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$$

GD update:

$$m_{new} = m_{old} - \alpha\frac{\partial J}{\partial m}$$

$$b_{new} = b_{old} - \alpha\frac{\partial J}{\partial b}$$

First Iteration

$m = -1,\ b = 1$

for (1,3): $\hat{y}_1 = (-1)\times 1 + 1 = 0$

for (3,6): $\hat{y}_2 = (-1)\times 3 + 1 = -2$

Errors: $3 - 0 = 3$

$6 - (-2) = 8$

Gradients: $\frac{\partial J}{\partial m} = -1(3\times 1 + 8\times 3) = -1\times(3+24) = -27$

$$\frac{\partial J}{\partial b} = -1(3+8) = -11$$

Update: $m_{new} = -1 + 2.7 = 1.7$

~~m old~~ =

$b_{new} = 1 + 1.1 = 2.7$