$$BCE = \begin{cases} y * -ln(\hat{y}), & \text{if } \hat{y} = 1 \\ (1-y) * -ln(1-\hat{y}), & \text{if } \hat{y} = 0 \end{cases}$$

When ŷ = 1:

$$\frac{J(\hat{y}+\epsilon, y) - J(\hat{y}, y)}{(\hat{y}+\epsilon) - \hat{y}} = \frac{[y * -ln(\hat{y}+\epsilon)] - [y * -ln(\hat{y})]}{(\hat{y}-\hat{y}) + \epsilon} = \frac{-ln(\hat{y}+\epsilon) + -[-ln(\hat{y})]}{\epsilon}$$

$$= \frac{-ln(\hat{y}+\epsilon) + ln(\hat{y})}{\epsilon} = \frac{-ln(\hat{y}+\epsilon) + ln(\hat{y})}{\epsilon} * \frac{-1}{-1} = \frac{ln(\hat{y}+\epsilon) - ln(\hat{y})}{-\epsilon}$$

$$= \frac{ln(\frac{\hat{y}+\epsilon}{\hat{y}})}{-\epsilon} = -\frac{1}{\epsilon} * ln(\frac{\hat{y}+\epsilon}{\hat{y}}) = -\frac{1}{\epsilon} * ln(\frac{\hat{y}}{\hat{y}} + \frac{\epsilon}{\hat{y}}) = -\frac{1}{\epsilon} * ln(1 + \frac{\epsilon}{\hat{y}})$$

$$= -ln([1 + \frac{\epsilon}{\hat{y}}]^{\frac{1}{\epsilon}}) \;\; ; \;\; n = \frac{\epsilon}{\hat{y}} \;\; ; \;\; \frac{n}{\epsilon} = \frac{1}{\hat{y}} \;\; ; \;\; \frac{n}{\epsilon} * \frac{1}{n} = \frac{1}{\hat{y}} * \frac{1}{n} \;\; ; \;\; \frac{1}{\epsilon} = \frac{1}{\hat{y}} * \frac{1}{n}$$

$$-ln([1+n]^{\frac{1}{\hat{y}} * \frac{1}{n}}) = -ln([(1+n)^{\frac{1}{n}}]^{\frac{1}{\hat{y}}}) = -\frac{1}{\hat{y}} ln([1+n]^{\frac{1}{n}}) = -\frac{1}{\hat{y}} ln(e) = -\frac{1}{\hat{y}} * 1 = -\frac{1}{\hat{y}}$$

When ŷ = 0:

$$\frac{J(\hat{y}+\epsilon, y) - J(\hat{y}, y)}{(\hat{y}+\epsilon) - \hat{y}} = \frac{[(1-y) * -ln(1 - [\hat{y}+\epsilon])] - [(1-y) * -ln(1-\hat{y})]}{(\hat{y}-\hat{y}) + \epsilon}$$

$$= \frac{[-ln(1 - [\hat{y}+\epsilon])] - [-ln(1-\hat{y})]}{\epsilon} = \frac{-ln(1-\hat{y}-\epsilon) + ln(1-\hat{y})}{\epsilon}$$

$$= \frac{-ln(1-\hat{y}-\epsilon) + ln(1-\hat{y})}{\epsilon} * \frac{-1}{-1} = \frac{ln(1-\hat{y}-\epsilon) - ln(1-\hat{y})}{-\epsilon} = \frac{ln(\frac{1-\hat{y}-\epsilon}{1-\hat{y}})}{-\epsilon} = -\frac{1}{\epsilon} * ln(\frac{1-\hat{y}-\epsilon}{1-\hat{y}})$$

$$= -\frac{1}{\epsilon} * ln(\frac{1}{1-\hat{y}} - \frac{\hat{y}}{1-\hat{y}} - \frac{\epsilon}{1-\hat{y}}) = -\frac{1}{\epsilon} * ln(\frac{1-\hat{y}}{1-\hat{y}} - \frac{\epsilon}{1-\hat{y}}) = -\frac{1}{\epsilon} * ln(1 - \frac{\epsilon}{1-\hat{y}})$$

$$= -ln([1 + \frac{-\epsilon}{1-\hat{y}}]^{\frac{1}{\epsilon}}) \;\; ; \;\; n = \frac{-\epsilon}{1-\hat{y}} \;\; ; \;\; \frac{n}{\epsilon} = \frac{-1}{1-\hat{y}} \;\; ; \;\; \frac{n}{\epsilon} * \frac{1}{n} = -\frac{1}{1-\hat{y}} * \frac{1}{n} \;\; ; \;\; \frac{1}{\epsilon} = (-\frac{1}{1-\hat{y}}) * \frac{1}{n}$$

$$-ln([1+n]^{(-\frac{1}{1-\hat{y}}) * \frac{1}{n}}) = -ln([(1+n)^{\frac{1}{n}}]^{\frac{-1}{1-\hat{y}}}) = -\frac{-1}{1-\hat{y}} ln([1+n]^{\frac{1}{n}}) = \frac{1}{1-\hat{y}} ln(e) = \frac{1}{1-\hat{y}} * 1 = \frac{1}{1-\hat{y}}$$