

1 - Price

$$\ln(q(y)) = \ln\left(\frac{1}{\sqrt{2\pi}\sigma^2} \cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right) = \quad (1)$$

$$\ln\left(\frac{1}{\sqrt{2\pi}\sigma^2}\right) + \ln\left(e^{-\frac{(y-\mu)^2}{2\sigma^2}}\right) =$$

$$\ln(1) - \ln(\sqrt{2\pi}) - \ln(\sqrt{\sigma^2}) - \frac{(y-\mu)^2}{2\sigma^2} =$$

$$e_\sigma = \frac{\partial \ln(q(y))}{\partial \sigma} = -\frac{1}{\sigma} - \frac{(y-\mu)^2}{2} \cdot \left(-\frac{1}{\sigma^3}\right) \cdot 2\sigma =$$

$$-\frac{1}{\sigma} + \frac{(y-\mu)^2}{\sigma^3}$$

$$\Delta d = \pi \cdot R \cdot e_\sigma = \pi \cdot R \cdot \left(\frac{(y-\mu)^2}{\sigma^3} - \frac{1}{\sigma}\right)$$

~~הערות: הנוסחה למעלה היא נכונה רק עבור התפלגות נורמלית. במקרה של התפלגות לוג-נורמלית, יש להשתמש בנוסחה שונה. בנוסף, יש לזכור שהנוסחה למעלה היא עבור התפלגות נורמלית, ולכן היא איננה תקפה עבור התפלגות לוג-נורמלית.~~

הערות: הנוסחה למעלה היא נכונה רק עבור התפלגות נורמלית. במקרה של התפלגות לוג-נורמלית, יש להשתמש בנוסחה שונה. בנוסף, יש לזכור שהנוסחה למעלה היא עבור התפלגות נורמלית, ולכן היא איננה תקפה עבור התפלגות לוג-נורמלית.

$$E[\Delta \sigma] = n \cdot \frac{d}{d\sigma} E[R]$$

(2)

$$E[R] = E[2my - y^2 - m^2] = 2m \cdot E[y] - E[y^2] - m^2$$

$$E[y] = \mu$$

$$E[y^2] = \text{Var}(y) + E(y)^2 = \sigma^2 + \mu^2$$

$$E[R] = 2m\mu - \sigma^2 - \mu^2 - m^2$$

$$\frac{d}{d\sigma} E[R] = -2\sigma$$

$$E[\Delta \sigma] = -2\sigma$$

הרכיב $E[\Delta \sigma]$ מושג כאשר $\sigma = 0$: $E[\Delta \sigma] = -2\sigma = 0$

$$\sigma = 0$$

אכן פשוט במובן זה הרכיב $E[\Delta \sigma]$ הוא ימין כוללת, ואין
הוא מציג הרצוא תואם.

2 > 1/e

$$e_i(y=0) = \frac{\partial \ln(P(y=0))}{\partial w_i} = \frac{\partial}{\partial w_i} \ln\left(\frac{1}{e^{w^T x_i} + 1}\right) = \textcircled{1}$$

$$= -\frac{1}{e^{w^T x_i} + 1} \cdot e^{w^T x_i} \cdot x_i = -\frac{1}{1 + e^{-w^T x_i}} \cdot x_i = -P(y=1) x_i$$

$$e_i(y=1) = \frac{\partial \ln(P(y=1))}{\partial w_i} = \frac{\partial}{\partial w_i} \ln(1 + e^{-w^T x_i})^{-1} =$$

$$= -\frac{1}{1 + e^{-w^T x_i}} \cdot e^{-w^T x_i} \cdot (-x_i) = P(y=0) x_i = [1 - P(y=1)] x_i$$

~~$e_i(y=0)$~~ $e_i(y=y) = [y - P(y=1)] \cdot x_i$ ②

$y \in \{0, 1\}$

$$\Delta w_i = n \cdot R \cdot e_i = n \cdot R [y - P(y=1)] x_i \quad \textcircled{3}$$

אם נניח n התצפיות תחתיות מהלך לפני כל עדכון
מקומי (קומה) שבתוך $x^T w$ כח R (אם R הוא 1)
→ R הוא מקומי (אם R הוא 1) ויש R הוא 1 בתוספת
שם כח R (אם R הוא 1) ויש R הוא 1 בתוספת

אם n הוא מספר התצפיות n הוא מספר התצפיות
הוא (אם n הוא מספר התצפיות n הוא מספר התצפיות)
הוא (אם n הוא מספר התצפיות n הוא מספר התצפיות)
הוא (אם n הוא מספר התצפיות n הוא מספר התצפיות)