№ 1

$$E_{x,y} E_{X^\ell} (y - a_{X^\ell}(x))^2 = E_{x,y} E_{X^\ell} (y - E(y|x) + E(y|x) - a_{X^\ell}(x))^2 = E_{x,y} \overbrace{E_{X^\ell} (y - E(y|x))^2}^{\text{НЕ ЗАВИСИТ ОТ } X^\ell} + E_{x,y} E_{X^\ell} (E(y|x) - E_{X^\ell} a_{X^\ell} + E_{X^\ell} a_{X^\ell} - a_{X^\ell}(x))^2 + \underbrace{2 E_{x,y} E_{X^\ell} (y - E(y|x)) \cdot (E(y|x) - a_{X^\ell}(x))}_{=:A} =$$

$$A = 2 E_{x,y} E_{X^\ell} \underbrace{(y - E(y|x))}_{\text{не зав от } E_{X^\ell}} \cdot (E(y|x) - a_{X^\ell}(x)) = 2 E_{x,y} \underbrace{(y - E(y|x))}_{=0} \cdot E_{X^\ell} (E(y|x) - a_{X^\ell}(x)) = 0$$

$$= E_{x,y} (y - E(y|x))^2 + E_{x,y} E_{X^\ell} (E(y|x) - E_{X^\ell} a_{X^\ell})^2 + E_{x,y} E_{X^\ell} (a_{X^\ell} - E_{X^\ell} a_{X^\ell})^2 + 2 E_{x,y} E_{X^\ell} (E(y|x) - E_{X^\ell} a_{X^\ell}) \cdot (E_{X^\ell} a_{X^\ell} - a_{X^\ell}(x)) = \text{noise} + \text{bias} + \text{variance} + 2R,$$ где:

$$R = E_{x,y} E_{X^\ell} (E(y|x) - E_{X^\ell} a_{X^\ell}) \underbrace{(E_{X^\ell} a_{X^\ell} - a_{X^\ell}(x))}_{\to 0} = 0$$

И т.д.

№2

$$a(x) = \frac{1}{N} \sum_{m=1}^{N} a_m(x)$$

Bias: $E_{x,y}\left(\left(E_{X^\ell}\left(\frac{1}{N} \sum a_{X^\ell}(x)\right) - E(y|x)\right)^2\right) =$

$$= E_{x,y}\left[\left[\frac{1}{N} \sum_{m=1}^{N} E_{X^\ell}\left[a_{X^\ell}(x) - E(y|x)\right]\right]^2\right] =$$

$$= E_{x,y}\left[\left(E_{X^\ell}\left(a_{X^\ell}(x) - E(y|x)\right)\right)^2\right] =$$

$$= E_{x,y}\left[\left[E_{X^\ell}\left[a_{X^\ell}(x)\right] - E(y|x)\right]^2\right] \Rightarrow$$

$\Rightarrow$ Bias тот же, что у одного алгоритма

Variance:

$$E_{x,y}\left[E_{X^\ell}\left[\underbrace{\left(\frac{1}{N} \sum a_{X^\ell}(x) - E_{X^\ell}\left[\frac{1}{N} \sum_{n=1}^{N} a_{X^\ell}(x)\right]\right)^2}_{A}\right]\right]$$

- $A = \frac{1}{N^2}\left(\sum_{n=1}^{N}\left[a_{X^\ell}(x) - E_{X^\ell}\left[a_{X^\ell}(x)\right]\right]\right)^2 =$

$$= \frac{1}{N^2}\left\{\sum_{n=1}^{N}\left(a_{X^\ell}(x) - E_{X^\ell}\left(a_{X^\ell}(x)\right)\right)^2 +\right.$$

$$+ \frac{1}{N^2} \sum_{n \neq k} \underbrace{\left[a_n{}_{X^\ell}(x) - E_{X^\ell}\left(a_{X^\ell}(x)\right)\right] \cdot \left[a_k{}_{X^\ell}(x) - E_{X^\ell}\left(a_k{}_{X^\ell}(x)\right)\right]}_{B_n \cdot B_k}$$

$\Downarrow$

- $E_{x,y} E_{X^\ell}(A) = \frac{1}{N^2} E_{x,y}\left[E_{X^\ell}\left[\sum_{n \neq k} B_n \cdot B_k\right]\right] +$

$$+ \frac{1}{N^2} E_{x,y}\left[E_{X^\ell}\left(\sum_{n=1}^{N} B_n^2\right)\right] \ominus$$

$\Rightarrow \frac{1}{N} E_{x,y}\left[E_{x^c} B_1^2\right] + \frac{N(N-1)}{N^2} E_{x,y}\left[E_{x^c} B_1 \cdot B_2\right] =$

$= \frac{1}{N}$ base_bias $+ \frac{N-1}{N} \cdot r \cdot b$, где

base_bias - bias базового алг-ма;

r - коэф. кор-ции двух баз. алг-ов (у всех один.)

$b := var^2(a_{x^c}(x))$

Если алг-мы некор-ны, то в разброс будет в N раз меньше

№ 3

$r = \frac{cov(\xi, \eta)}{\sigma_\xi \cdot \sigma_\eta}$. Пусть $D(x_j, x_j) = \sigma$

$r(x_i, x_j) = \rho_i = \frac{cov(x_i, x_j)}{\sigma^2}$

$D(\overline{X}) = \frac{1}{N^2} E\left(\sum x_i\right)^2 - \frac{1}{N^2}\left(E \sum x_i\right)^2 =$

$= \frac{1}{N^2}\left(\sum_{i=1}^{N} E x_i^2 + \cancel{2} \sum_{i \neq j} E x_i x_j - \sum_{i=1}^{N}(E x_i)^2 - \cancel{2} \sum_{i \neq j} E x_i E x_j\right.$

$= \frac{1}{N^2}\left(N \cdot E X_1^2 + N(N-1) E X_1 X_2 - N \cdot (E X_1)^2 - N(N-1) E X_1 E X_2\right.$

$= \frac{1}{N^2}\left(N\left(E X_1^2 - (E X_1)^2\right) + N^2\left(E X_1 X_2 - E X_1 E X_2\right) -\right.$

$\left.- N\left(E X_1 X_2 - E X_1 E X_2\right)\right) =$

$= \frac{\sigma^2}{N} + \rho\sigma^2 - \frac{1}{N}\rho\sigma^2 = \rho\sigma^2 + (1-\rho)\frac{\sigma^2}{N}$

ч.т.д.