Макс и Нюша хотят оценить три неизвестные константы, μ_1, μ_2 и μ_3. К несчастью, у них всего три независимых наблюдения: $X_1 \sim \mathcal{N}(\mu_1, 1)$, $X_2 \sim \mathcal{N}(\mu_2, 1)$ и $X_3 \sim \mathcal{N}(\mu_3, 1)$.

Поскольку оценить важно все три константы, Макс и Нюша хотят, чтобы оценки минимизировали общую величину риска $R = \mathrm{E}((\hat{\mu}_1 - \mu_1)^2 + (\hat{\mu}_2 - \mu_2)^2 + (\hat{\mu}_3 - \mu_3)^2)$.

Максим использует метод максимального правдоподобия для получения вектор-столбца оценок $\hat{\mu}^{ML}$. А Нюша просто любит число ноль, поэтому использует оценку $\hat{\mu}^N = \alpha \cdot 0 + (1-\alpha)\hat{\mu}^{ML}$, где $\alpha = 1/X'X$ и X — вектор-столбец из X_i.

1. Несут ли величины X_2 и X_3 информацию о параметре μ_1?
2. Какие оценки получит Макс?
3. Чему будет равна общая величина риска для оценок Макса?
4. Какую формулу использует Нюша для оценки μ_1?
5. У кого общая величина риска будет выше, у Макса или у Нюши?

Без доказательства можно пользоваться утверждением:

Если $g(X)$ — дифференциируемая функция $g : \mathbb{R}^n \to \mathbb{R}$ и $X \sim \mathcal{N}(\mu, \sigma^2 I)$, где I — единичная матрица, то

$$\mathrm{E}((X_i - \mu)g(X)) = \mathrm{E}\left(\frac{\partial g(X)}{\partial X_i}\right).$$

Ответы:

1. нет
2. $\hat{\mu}_i = X_i$
3. $R^{ML} = 3$
4. $\hat{\mu}_1 = \left(1 - \frac{1}{X_1^2 + X_2^2 + X_3^2}\right) X_1$
5. У Нюши.

   $\hat{\mu}^N = \alpha \cdot 0 + (1-\alpha)X = X - g(X)$

   $$R^N = \mathrm{E}[(X - g(X) - \mu)^T(X - g(X) - \mu)] = \mathrm{E}[(X-\mu)^T(X-\mu)] + \mathrm{E}(g(X)^T g(X)) - 2\,\mathrm{E}[(X-\mu)^T g(X)]$$

   Замечаем, что $\mathrm{E}[(X-\mu)^T(X-\mu)] = 3$.

   Из предложенной теоремы следует, что $\mathrm{E}[(X-\mu)^T g(X)] = \sum_i \mathrm{E}\left(\frac{\partial g_i(X)}{\partial X_i}\right) = \mathrm{E}(g(X)^T g(X))$.

   Следовательно,

   $$R^N = 3 + \mathrm{E}(g(X)^T g(X)) - 2\,\mathrm{E}(g(X)^T g(X)) = 3 - \mathrm{E}(g(X)^T g(X)) < 3$$