1 4.1

Согласно теореме байеса
$P(y|x) = \frac{P(x|y)P(y)}{P(x)} = \frac{\prod P(x^{(k)}|y)P(y)}{P(x)}$
Предсказание байесовского классификатора на объекте x вычисляется следующим образом:
$\hat{y} = argmax_y P(y|x) = argmax_y \frac{\prod P(x^{(k)}|y)P(y)}{P(x)}$

$P(x) = const \ \forall y$
$P(y) = const \ \forall y$

$argmax_y P(y|x) = argmax_y P(x|y) = argmax_y \prod P(x^{(i)}|y) =$
$= argmax \frac{1}{(2\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum(x^{(k)}-\mu_{yk})^2} = argmin_y(\sum(x^k - \mu_{yk})^2)$

Таким образом, классификатор выберет класс, расстояние от центра которого до объекта минимально

2 4.2

В случае не вероятностной классификации ROC-кривая принимает вид треугольника с вершинами в (0, 0), (1, 1) и (FPR, TPR).
Пусть доля класса 1 в обучающей выборке равна k, а классификатор возвращает случайно 1 с вероятностью p, 0 с вероятностью $1 - p$.
$E(TPR) = \frac{p*kn}{kn} = p$
$E(FPR) = \frac{p*(1-k)n}{(1-k)n} = p$

Таким образом, вне зависимости от параметров p и k мат.ожидание TPR и FPR совпадают, а следовательно, точка (FPR, TPR) в среднем лежит на диагонали квадрата с вершинами (0, 0) и (1, 1). Площаль такой ROC-кривой равна 1/2, что и требовалось показать.

3 4.3

$E_N = P(y \neq y_n)$
При $n \to \infty$ $x_n \to x$ из предположения о равномерности объектов в пространстве
Так как условные вероятности непрерывны по x:
$P(y \neq y_n) = P(1|x)P(0|x_n) + P(0|x)P(1|x_n) \to 2P(1|x)P(0|x)$
$2P(1|x)P(0|x) \leq 2min\{P(0|x)P(1|x)\}$
Таким образом, в пределе
$E_N \leq 2E_B$