Домашнее задание №1.

35.

$L(y', y) = (y'-y)^2$ - ф. п., её min значение достигается при $y'=y$

Теперь перейдём к $f^*(x)$

$$f^*(x) = \arg\min_c E((Y-c)^2 | X=x) =$$

$$= \arg\min_c E(Y^2 - 2Yc + c^2 | X=x) =$$

$$= \arg\min_c \{ E(Y^2|X=x) - 2E(Yc|X=x) + c^2 E \} \Rightarrow$$

Первое и второе слагаемые не зависят от c, а значит min будет достигаться при
$c = E(Y|X=x)$
собственно в итоге:

$$f^*(x) = E(Y|X=x)$$

Тогда $R(f^*)$ = по определ. ф. о.

$$= E((Y - f^*(x))^2) = E((Y - E(Y|X=x))^2) =$$

А это выражение не что иное как дисперсия условного события

$$= R(f^*) = Var(Y|X)$$

3c.

$L(y', y) = |y' - y|$

$$R(d) = \int_X \int_Y L(d(x), y)\, P(y|x)\, dy\, p(x)\, dx$$

$$\int_Y L(d(x), y)\, p(y|x)\, dy = E(L(d(X), Y) / X = x) \Rightarrow$$

$$\Rightarrow E(|\overset{=c}{d(x)} - Y| \mid X = x) \to \min_c$$

- $\text{median}(Y / X = x) = m$

  $P(Y < m | X = x) = P(Y > m | X = x) = 1/2$ или

  $\int_{Y<m} p(y|x)\,dy = \int_{Y>m} p(y|x)\,dy = 1/2$

- Пусть $a > 0$ и рассмотрим следующее

$$E(|m + a - Y| | X = x) = \int_Y |m + a - y|\, p(y|x)\, dy =$$

$$= \underbrace{\int_{Y<m+a} (m + a - y)\, p(y|x)\, dy}_{1} + \underbrace{\int_{Y>m+a} (y - m - a)\, p(y|x)\, dy}_{2}$$

1 $\int_{Y<m} (\dots)\, p\, dy + \int_{m<Y<m+a} (\dots)\, p\, dy$

2 $\int_{Y>m} (\dots)\, p\, dy - \int_{m<Y<m+a} (\dots)\, p\, dy$

Подставим в выражение выше получим.

$$\int_{Y>m} (y - m)\, p\, dy + \int_{Y<m} (m - y)\, p\, dy + a\Big(\overbrace{\int_{Y<m} p\, dy}^{=1/2} - \overbrace{\int_{Y>m} p\, dy}^{=1/2}\Big) +$$

$$+ 2\int_{m<Y<m+a} (m + \overset{>0}{a} - y)\, \overset{>0}{p}\, dy = \underbrace{\int_Y |y - m|\, p\, dy}_{E(|Y - m| / X = x)} + 2\int_{m<Y<m+a} (m + a - y)\, p\, dy$$

И чтобы уменьшить выражение берём $a = 0$

- Аналогичные рассуждения при $a < 0$

А значит вывод один $c = \text{median}(Y | X = x)$ и при ней достигается min.

37.

Предположим, что ф. п. должна иметь такой вид

$$L(y', y) = \begin{cases} 0, & \text{если } y' = y \\ 1, & \text{если } y' \neq y \end{cases}$$

$$R(f) = \int_{X \times Y} L(f(x), y)\, p(x, y)\, dx dy$$

далее удобно будет взять приближенное значение

$$\approx \tilde{R}(f) = 1/N \sum_{i=1}^{N} L(f(x^i), y^i)$$

* Если брать $f(x)$ = моде, то сумма будет из 0 и 1, но этих самых единиц будет меньше чем при других $f(x)$, т.к. изначально выбрали моду - наиболее частое в выб. число.

38.

Дополним задачу.

$$Pr(y=0 \mid X=x) = \left\{ \frac{Pr(X=x \mid y=0)\, Pr(y=0)}{Pr(X=x)} \right\}$$

$$Pr(y=1 \mid X=x) = \left\{ \frac{Pr(X=x \mid y=1)\, Pr(y=1)}{Pr(X=x)} \right\}$$

$$\Rightarrow R(f^*) = \sum_{y \in \{0,1\}} l_y \, Pr(Y=y \mid X=x)$$

$$f^*(x) = \arg\min_{y \in \{0,1\}} R(f(x)) = \arg\min_{y \in \{0,1\}} \sum_{y \in Y} l_y \, Pr(Y=y \mid X=x)$$

$$\Rightarrow \begin{cases} l_0 < l_1 \Rightarrow R \text{ будет min при } y=1 \\ l_0 > l_1 \Rightarrow R \text{ будет min при } y=0 \\ l_0 = l_1 \Rightarrow \text{выбор можно любой.} \end{cases}$$

Ну и в основе всего этого доказано что

$$f^*(x) = \arg\min_{y} l_y \, Pr(Y=y \mid X=x).$$

39.

Необходимо вывести Байесов классификатор.

$$f^*(x) = \arg\max_{y \in Y} Pr(y|x) = \arg\max_{y \in Y} \frac{P(x|y)\, Pr(y)}{P(x)} =$$

$$= \arg\max_{y \in Y} P(x|y)\, Pr(y)$$

* Исключили $P(x)$ т.к. он не зависит от y

$Pr(y=k)$ – апост. вероят. что объект класса k

$Pr(x|y=k)$ для объекта x классификатор выбирает класс k с min ожид. риском.

$$R(f) = \sum_{y'=1}^{K} L(y', y)\, Pr(y|X)$$

$$f(x) = \arg\min_{k} \sum_{y'=1}^{k} L(y', \overset{k}{y})\, Pr(y'|X) =$$

$$= \arg\min_{k} \sum_{y=1}^{k} L(y', \overset{k}{y})\, Pr(y'|X),$$

$$Pr(y'|X) = \frac{Pr(X|y'=k)\, Pr(y'=k)}{\sum_{j=1}^{k} Pr(X|y=j)\, Pr(y'=j)}$$

класс k min ожид. риск:

$$k = \arg\min_{c = \{1, k\}} \sum_{j=1}^{k} L(c, j)\, Pr(y'=j\,|X=x)$$