Oulun yliopisto
Matemaattisten tieteiden yksikkö
Tilastollinen hahmontunnistus
Harjoitus 6

1. Todista, että kaikkien luokittelijoiden joukossa Bayesin minimiriskiluokittimella on pienin riski (luentojen lause 3.14). (Vihje: Eräs tapa päästä alkuun on havaita, että luokittimen g riski $R(g) = \mathbb{E}\lambda(J, g(X))$ on diskreetin satunnaismuuttujan odotusarvo ja siis helposti laskettavissa; esim. Tuominen: Todennäköisyyslaskenta I, Lause 3.1.8 (i).)

2. Osoita, että Bayesin luokitin on minimiriskiluokitin, joka saadaan, kun tappioksi valitaan

$$\lambda(j,k) = \begin{cases} 0, & \text{kun } j = k, \\ 1, & \text{kun } j \neq k. \end{cases}$$

3. Olkoot kahden luokan tapauksessa luokkien 1 ja 2 luokkatiheysfunktiot f_1 ja f_2 seuraavat

$$f_1(x) = \begin{cases} \frac{1}{2} - \frac{1}{2}x, & \text{kun} -1 \leq x \leq 1, \\ 0 & \text{muuten,} \end{cases} \qquad f_2(x) = \begin{cases} \frac{1}{2} + \frac{1}{2}x, & \text{kun} -1 \leq x \leq 1, \\ 0 & \text{muuten.} \end{cases}$$

Luokan 1 prioritodennäköisyys on p, $0 < p < 1$ (ja siis luokan 2 vastaavasti $1 - p$).

Mitkä ovat luokan 1 prioritodennäköisyyden arvoa p vastaavan Bayesin luokittimen g_p^* päätösalueet $A_1^*(p), A_2^*(p)$ sekä sen ehdolliset virhetodennäköisyydet

$$e_j(g_p^*) = P(g_p^*(X) \neq j | J = j), \quad j = 1, 2?$$

4. Olkoot kahden luokan tapauksessa luokkien 1 ja 2 luokkatiheysfunktiot f_1 ja f_2 samat kuin tehtävässä 3 ja $p = 1/2$. Tarkastellaan minimiriskiluokitinta hylkäysmahdollisuudella, kun hylkäyksestä tuleva tappio on t. Millä välillä sijaitsevat t:n järkevät arvot? Olkoon g_t^* kyseinen minimiriskiluokitin ja, kuten luennoilla, olkoon $r(t)$ todennäköisyys, että g_t^* hylkää ja $e(t)$ todennäköisyys, että g_t^* ei hylkää mutta tekee virheen. Hahmottele virhe-hylkäyskäyrä $t \mapsto (e(t), r(t))$.

5. Olkoot kahden luokan tapauksessa luokkien 1 ja 2 luokkatiheysfunktiot f_1 ja f_2 samat kuin tehtävässä 3. Hahmottele Neymanin-Pearsonin luokittimen

$$\frac{f_1(x)}{f_2(x)} \underset{2}{\overset{1}{\gtrless}} u$$

toimintaominaiskäyrä $u \mapsto (e_2(u), 1 - e_1(u))$, $u \geq 0$, missä $e_j(u)$ on luokittimen virhe luokassa j, $j = 1, 2$.