On Discriminative Learning of Prediction Uncertainty

Vojtech Franc   Daniel Prusa

A. Supplementary materials

A.1. Proof of Theorem 1

Task 1

$$\max_{h,c} \phi(c) \quad \textit{s.t.} \quad R_S(h,c) \leq \lambda\,, \tag{1}$$

Theorem 1 *Let (h,c) be an optimal solution to (1). Then, (h_B,c), where h_B is the optimal Bayes classifier, is also optimal to (1).*

Proof 1 *It is sufficient to show that (h_B,c) is feasible to (1), i.e., that $R_S(h_B,c) \leq \lambda$. Then (h_B,c) attains the maximum objective value $\phi(c)$. Derive*

$$\begin{aligned}
R_S(h_B,c) &= \frac{1}{\phi(c)} \int\limits_{\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y)\,\ell(y,h_B(x))\,c(x)\,dx \\
&= \frac{1}{\phi(c)} \int\limits_{\mathcal{X}} p(x)c(x) \left(\sum_{y\in\mathcal{Y}} p(y\,|\,x)\,\ell(y,h_B(x))\right) dx \\
&\leq \frac{1}{\phi(c)} \int\limits_{\mathcal{X}} p(x)c(x) \left(\sum_{y\in\mathcal{Y}} p(y\,|\,x)\,\ell(y,h(x))\right) dx \\
&= \frac{1}{\phi(c)} \int\limits_{\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y)\,\ell(y,h(x))\,c(x)\,dx \\
&= R_S(h,c) \leq \lambda.
\end{aligned}$$

□

A.2. Proof of Theorem 2

The presented proof of the theorem uses Lemmas 2 and 3, both derived based on Lemma 1 bellow.

Lemma 1 *Let $f : \mathcal{X} \to \mathbb{R}_+$[1] and $g : \mathcal{X} \to \mathbb{R}$ be measurable functions such that $\int_{\mathcal{X}} f(x)dx > 0$ and $g(x) > 0$ for all $x \in \mathcal{X}$. Then it holds $\int_{\mathcal{X}} g(x)f(x)dx > 0$.*

Proof 2 *For $n \in \mathbb{N}_+$, define functions*

$$f_n(x) = \begin{cases} f(x) & \textit{if } g(x) \geq \frac{1}{n}, \\ 0 & \textit{otherwise.} \end{cases}$$

The sequence $\{f_n\}_{n=1}^{\infty}$ is monotone and converges to f. Using the monotone convergence theorem (Stein & Shakarchi, 2009), derive

$$0 < \int\limits_{\mathcal{X}} f(x)dx = \lim_{n\to\infty} \int\limits_{\mathcal{X}} f_n(x)dx,$$

[1] We use $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{N}_+$ to denote the set of real numbers, non-negative real numbers and positive integers, respectively.

which implies there is a $k \in \mathbb{N}_+$ *such that* $\int_{\mathcal{X}} f_k(x)dx > 0$*, hence we conclude*

$$\int_{\mathcal{X}} g(x)f(x)dx \geq \int_{\mathcal{X}} g(x)f_k(x)dx \geq \int_{\mathcal{X}} \frac{1}{k} f_k(x)dx > 0.$$

□

Lemma 2 *Let* $f : \mathcal{X} \to \mathbb{R}_+$ *and* $g : \mathcal{X} \to \mathbb{R}$ *be measurable functions such that* $\int_{\mathcal{X}} f(x)dx > 0$ *and* $g(x) > b$ *for all* $x \in \mathcal{X}$ *and some* $b \in \mathbb{R}$*. Then it holds* $\int_{\mathcal{X}} g(x)f(x)dx > b \int_{\mathcal{X}} f(x)dx$*.*

Proof 3 *By Lemma 1, we have*

$$\int_{\mathcal{X}} (g(x) - b)f(x)dx > 0,$$

thus

$$\int_{\mathcal{X}} g(x)f(x)dx = \int_{\mathcal{X}} (g(x) - b)f(x)dx + \int_{\mathcal{X}} bf(x)dx > b \int_{\mathcal{X}} f(x)dx.$$

□

Lemma 3 *Let* $f : \mathcal{X} \to \mathbb{R}_+$ *and* $g : \mathcal{X} \to \mathbb{R}$ *be measurable functions such that* $\int_{\mathcal{X}} g(x)f(x)dx > 0$ *and* $g(x) < 1$ *for all* $x \in \mathcal{X}$*. Then it holds* $\int_{\mathcal{X}} f(x)dx > \int_{\mathcal{X}} g(x)f(x)dx$*.*

Proof 4 $\int_{\mathcal{X}} g(x)f(x)dx > 0$ *implies* $\int_{\mathcal{X}} f(x)dx > 0$*. Since it holds* $\forall x \in \mathcal{X} : (1 - g(x)) > 0$*, Lemma 1 yields*

$$0 < \int_{\mathcal{X}} (1 - g(x))f(x)dx = \int_{\mathcal{X}} f(x)dx - \int_{\mathcal{X}} g(x)f(x)dx,$$

which implies $\int_{\mathcal{X}} f(x)dx > \int_{\mathcal{X}} g(x)f(x)dx$*.* □

Task 2

$$\max_{c \in [0,1]^{\mathcal{X}}} \int_{\mathcal{X}} p(x)c(x)dx \quad \text{s.t.} \quad \int_{\mathcal{X}} p(x)c(x)\overline{r}(x)dx \leq 0\,. \tag{2}$$

Theorem 2 *A selection function* $c^* : \mathcal{X} \to [0, 1]$ *is an optimal solution to (2) if and only if it holds*

$$\int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)dx = \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)dx, \tag{3}$$

$$\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx = \begin{cases} -\frac{\rho(\mathcal{X}_{\overline{r}(x)<b})}{b} & \text{if} \quad b > 0\,, \\ \int_{\mathcal{X}_{\overline{r}(x)=0}} p(x)dx & \text{if} \quad b = 0\,, \end{cases} \tag{4}$$

$$\int_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c^*(x)dx = 0 \tag{5}$$

where

$$\rho(\mathcal{X}') = \int_{\mathcal{X}'} p(x)\overline{r}(x)\,dx \tag{6}$$

is the conditional expectation of $\overline{r}(x)$ *over inputs in* $\mathcal{X}' \subseteq \mathcal{X}$*, and*

$$b = \sup\{a \mid \rho(\mathcal{X}_{\overline{r}(x)\leq a}) \leq 0\} \geq 0\,. \tag{7}$$

Proof 5 *Let* $F(c) = \int_{\mathcal{X}} p(x)c(x)dx$ *denote the objective function of (2). Observe that* $b \geq 0$*, because* $\rho(\mathcal{X}_{\overline{r}(x)\leq 0}) \leq 0$*.*

Case 1 $b > 0$.
Claim I *Each* $c^* : \mathcal{X} \to [0,1]$ *which fulfils (3), (4) and (5) is feasible to (2) and*

$$F(c^*) = \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)dx - \frac{1}{b}\rho(\mathcal{X}_{\overline{r}(x)<b}). \tag{8}$$

Proof of Claim I.

Equality (8) is simply obtained by summing LHS and RHS of (3), (4) and (5). Next, verify the constraint of (2).

Observe that $\int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)(c^*(x)-1)dx \overset{(3)}{=} 0$ *implies* $\int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)(c^*(x)-1)\overline{r}(x)dx = 0$ *and*

$$\int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)\overline{r}(x)dx = \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)\overline{r}(x)dx \overset{(6)}{=} \rho(\mathcal{X}_{\overline{r}(x)<b}). \tag{9}$$

If $b < \infty$, *then*

$$\begin{aligned}\int\limits_{\mathcal{X}} p(x)c^*(x)\overline{r}(x)dx &\overset{(5)}{=} \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)\overline{r}(x)dx + \int\limits_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)\overline{r}(x)dx \\ &\overset{(9)}{=} \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)\overline{r}(x)dx + b\int\limits_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx \overset{(4),(6)}{=} \rho(\mathcal{X}_{\overline{r}(x)<b}) - \rho(\mathcal{X}_{\overline{r}(x)<b}) = 0.\end{aligned} \tag{10}$$

If $b = \infty$, *then*

$$\int\limits_{\mathcal{X}} p(x)c^*(x)\overline{r}(x)dx = \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)\overline{r}(x)dx \overset{(9)}{=} \rho(\mathcal{X}_{\overline{r}(x)<b}) \leq 0.$$

Claim II *Let* $c : \mathcal{X} \to [0,1]$ *be a feasible solution to (2) that violates at least one of the constraints (3), (4) and (5). Then,* $F(c) < F(c^*)$ *where* $c^* : \mathcal{X} \to [0,1]$ *is a confidence function fulfilling (3), (4), (5), and, w.l.o.g.,*

$$\forall x \in \mathcal{X}_{\overline{r}(x)<b} : c^*(x) = 1\,. \tag{11}$$

To prove Claim II, distinguish three cases.

Case 1.1 *Condition (5) is violated (observe that this is possible only if* $b < \infty$*), i.e.*

$$\int\limits_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx > 0. \tag{12}$$

Inequality (12) and Lemma 2 (applied to $f(x) = p(x)c(x)$ *and* $g(x) = \overline{r}(x)$*) yield*

$$\int\limits_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)\overline{r}(x)dx > b\int\limits_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx.$$

Hence, we can write

$$\int\limits_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)\overline{r}(x)dx = b'\int\limits_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx \tag{13}$$

for a suitable real number b' *such that*

$$b' > b > 0. \tag{14}$$

Based on the constraint of (2), derive

$$\int_{\mathcal{X}} p(x)c(x)\overline{r}(x)dx \overset{(13)}{=} \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)\overline{r}(x)dx + b\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx + b'\int_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx$$
$$\overset{(2)}{\leq} 0 \overset{(10)}{=} \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)\overline{r}(x)dx + b\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx. \tag{15}$$

Let $\sigma(x) = \frac{1}{b}\overline{r}(x)$. Inequality (15) can be rearranged and upper bounded as

$$\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx - \int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx + \frac{b'}{b}\int_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx \overset{(15)}{\leq} \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)(c^*(x)-c(x))\sigma(x)dx$$
$$\leq \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)dx - \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)dx \tag{16}$$

where the second inequality follows from $\forall x \in \mathcal{X}_{\overline{r}(x)<b} : \sigma(x) \leq 1$. From this we get

$$\int_{\mathcal{X}_{\overline{r}(x)\leq b}} p(x)c(x)dx \overset{(16)}{\leq} \int_{\mathcal{X}_{\overline{r}(x)\leq b}} p(x)c^*(x)dx - \frac{b'}{b}\int_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx\,. \tag{17}$$

Now, derive

$$F(c) = \int_{\mathcal{X}_{\overline{r}(x)\leq b}} p(x)c(x)dx + \int_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx \overset{(17)}{\leq} \int_{\mathcal{X}_{\overline{r}(x)\leq b}} p(x)c^*(x)dx - \left(\frac{b'}{b}-1\right)\int_{\mathcal{X}_{\overline{r}(x)>b}} p(x)c(x)dx$$
$$\overset{(12),(14)}{<} \int_{\mathcal{X}_{\overline{r}(x)\leq b}} p(x)c^*(x)dx = F(c^*).$$

Case 1.2 *Condition (5) holds, condition (4) is violated.*

If $\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx < -\frac{\rho(\mathcal{X}_{\overline{r}(x)<b})}{b}$, then obviously $F(c) < F(c^*)$. Hence, assume

$$\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx > -\frac{\rho(\mathcal{X}_{\overline{r}(x)<b})}{b}. \tag{18}$$

Analogically to (15), derive

$$\int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)\overline{r}(x)dx + b\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx \overset{(2)}{\leq} 0 \overset{(10)}{=} \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)\overline{r}(x)dx + b\int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx,$$

and

$$\int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)\sigma(x)dx + \int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx \leq \int_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)\sigma(x)dx + \int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx \tag{19}$$

where $\sigma(x) = \frac{1}{b}\overline{r}(x) < 1$ for all $x \in \mathcal{X}_{\overline{r}(x)<b}$.

Denote and derive

$$\Delta = \int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx - \int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx \overset{(4)}{=} \int_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx + \frac{\rho(\mathcal{X}_{\overline{r}(x)<b})}{b} \overset{(18)}{>} 0. \tag{20}$$

Then, (19) can be rewritten as

$$\int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)(c^*(x)-c(x))\sigma(x)dx \geq \Delta \overset{(20)}{>} 0. \tag{21}$$

Inequality (21) and Lemma 3 (applied to $g(x)=\sigma(x)<1$ *and* $f(x)=p(x)(c^*(x)-c(x)) \overset{(11)}{\geq} 0$ *over* $\mathcal{X}_{\overline{r}(x)<b}$*) yield*

$$\int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)(c^*(x)-c(x))dx > \Delta. \tag{22}$$

Now, combine and rearrange (20) and (22) to obtain

$$F(c^*) = \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)dx + \int\limits_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c^*(x)dx \overset{(22)}{>} \Delta \overset{(20)}{=} \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)dx + \int\limits_{\mathcal{X}_{\overline{r}(x)=b}} p(x)c(x)dx = F(c).$$

Case 1.3 *Conditions (4) and (5) hold, condition (3) is violated, i.e.*

$$\int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)dx < \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)dx. \tag{23}$$

Then,

$$F(c^*) = \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c^*(x)dx - \frac{\rho(\mathcal{X}_{\overline{r}(x)<b})}{b} \overset{(23)}{>} \int\limits_{\mathcal{X}_{\overline{r}(x)<b}} p(x)c(x)dx - \frac{\rho(\mathcal{X}_{\overline{r}(x)<b})}{b} = F(c).$$

Case 2 $b=0$.

This occurs only if $\int_{\mathcal{X}_{\overline{r}(x)<0}} p(x)\overline{r}(x)dx = 0$*. The constraint of (2) implies*

$$\int\limits_{\mathcal{X}_{\overline{r}(x)>0}} p(x)c(x)\overline{r}(x)dx = 0\,,$$

thus

$$\int\limits_{\mathcal{X}_{\overline{r}(x)>0}} p(x)c(x)dx = 0\,,$$

which confirms condition (5).

Finally, the obvious equations

$$\max_{c:\mathcal{X}\to[0,1]} \int\limits_{\mathcal{X}_{\overline{r}(x)<0}} p(x)c(x)dx = \int\limits_{\mathcal{X}_{\overline{r}(x)<0}} p(x)dx\,, \text{ and}$$
$$\max_{c:\mathcal{X}\to[0,1]} \int\limits_{\mathcal{X}_{\overline{r}(x)=0}} p(x)c(x)dx = \int\limits_{\mathcal{X}_{\overline{r}(x)=0}} p(x)dx$$

confirm condition (3) and (4), respectively.

□

A.3. Proof of Theorem 3

Task 3

$$\min_{s:\mathcal{X}\to\mathbb{R}} E(s) \tag{24}$$

$$\text{where} \quad E(s) = \int\limits_{\mathcal{X}} p(x)\,r(x) \int\limits_{\mathcal{X}} p(z)\,[\![s(x)\leq s(z)]\!]dz\,dx\,. \tag{25}$$

Remark 1 *For the sake of simplicity, for predicates $\varphi_1(x,z),\ldots,\varphi_k(x,z)$ and a function $f:\mathcal{X}\times\mathcal{X}\to\mathbb{R}$, we write*

$$\int\limits_{\mathcal{X}}\int\limits_{\substack{\varphi_1(x,z)\\ \vdots\\ \varphi_k(x,z)}} f(x,z)dz\,dx$$

to represent

$$\int\limits_{\mathcal{X}}\int\limits_{\mathcal{X}} f(x,z)[\![\varphi_1(x,z)\wedge\ldots\wedge\varphi_k(x,z)]\!]dz\,dx\,.$$

Theorem 3 *A function $s^*:\mathcal{X}\to\mathbb{R}$ is an optimal solution to $\min_{s:\mathcal{X}\to\mathbb{R}}E(s)$ if and only if*

$$\int\limits_{\mathcal{X}}\int\limits_{\substack{z\neq x\\ s^*(z)=s^*(x)}} \max\{r(x),r(z)\}p(x)p(z)dz\,dx=0 \tag{26}$$

and

$$\int\limits_{\mathcal{X}}\int\limits_{\substack{r(z)<r(x)\\ s^*(z)>s^*(x)}} (r(x)-r(z))\,p(x)p(z)dz\,dx=0. \tag{27}$$

Proof 6 *We first list four equalities to be used later, all of them easy to verify (note that $s:\mathcal{X}\to\mathbb{R}$ is any measurable function):*

$$\int\limits_{\mathcal{X}} r(x)p(x)\int\limits_{\substack{r(z)>r(x)\\ s(z)<s(x)}} p(z)dz\,dx=\int\limits_{\mathcal{X}} p(z)\int\limits_{\substack{r(x)<r(z)\\ s(x)>s(z)}} r(x)p(x)dx\,dz=\int\limits_{\mathcal{X}} p(x)\int\limits_{\substack{r(z)<r(x)\\ s(z)>s(x)}} r(z)p(z)dz\,dx\,, \tag{28}$$

$$\begin{aligned}\int\limits_{\mathcal{X}}\int\limits_{\substack{r(z)<r(x)\\ s(z)=s(x)}} r(x)p(x)p(z)dz\,dx=&\frac{1}{2}\int\limits_{\mathcal{X}}\int\limits_{\substack{z\neq x\\ s(z)=s(x)}} \max\{r(x),r(z)\}p(x)p(z)dz\,dx\\ &-\frac{1}{2}\int\limits_{\mathcal{X}}\int\limits_{\substack{z\neq x\\ r(z)=r(x)\\ s(z)=s(x)}} \max\{r(x),r(z)\}p(x)p(z)dz\,dx\,,\end{aligned} \tag{29}$$

$$\int\limits_{\mathcal{X}}\int\limits_{\substack{r(z)=r(x)\\ s(z)<s(x)}} r(x)p(x)p(z)dz\,dx=\frac{1}{2}\int\limits_{\mathcal{X}}\int\limits_{r(z)=r(x)} r(x)p(x)p(z)dz\,dx-\frac{1}{2}\int\limits_{\mathcal{X}}\int\limits_{\substack{r(z)=r(x)\\ s(z)=s(x)}} r(x)p(x)p(z)dz\,dx\,, \tag{30}$$

$$\int\limits_{\mathcal{X}}\int\limits_{\substack{r(z)=r(x)\\ s(z)=s(x)}} r(x)p(x)p(z)dz\,dx=2\int\limits_{\mathcal{X}}\int\limits_{\substack{z>x\\ r(z)=r(x)\\ s(z)=s(x)}} r(x)p(x)p(z)dz\,dx+\int\limits_{\mathcal{X}}\int\limits_{z=x} r(x)p(x)p(z)dz\,dx\,. \tag{31}$$

Since $\operatorname{argmin}_{s:\mathcal{X}\to\mathbb{R}}E(s)=\operatorname{argmin}_{s:\mathcal{X}\to\mathbb{R}}(E(s)-E(r))$*, it suffices to analyze minimizers of* $E(s)-E(r)$ *instead of*

$E(s)$. *Derive*

$$\begin{aligned}
E(s)-E(r) &= \int_{\mathcal{X}} \int_{s(z)\geq s(x)} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{r(z)\geq r(x)} r(x)p(x)p(z)dz\,dx \\
&= \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)\geq s(x)}} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{\substack{r(z)\geq r(x)\\ s(z)<s(x)}} r(x)p(x)p(z)dz\,dx \\
&= \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)>s(x)}} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{\substack{r(z)>r(x)\\ s(z)<s(x)}} r(x)p(x)p(z)dz\,dx \\
&+ \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)=s(x)}} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{\substack{r(z)=r(x)\\ s(z)<s(x)}} r(x)p(x)p(z)dz\,dx \\
&= F_1(s)+F_2(s)
\end{aligned}$$

where

$$\begin{aligned}
F_1(s) &= \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)>s(x)}} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{\substack{r(z)>r(x)\\ s(z)<s(x)}} r(x)p(x)p(z)dz\,dx \\
&\overset{(28)}{=} \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)>s(x)}} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)>s(x)}} r(z)p(x)p(z)dz\,dx \\
&= \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)>s(x)}} \left(r(x)-r(z)\right)p(x)p(z)dz\,dx
\end{aligned}$$

and

$$F_2(s) = \int_{\mathcal{X}} \int_{\substack{r(z)<r(x)\\ s(z)=s(x)}} r(x)p(x)p(z)dz\,dx - \int_{\mathcal{X}} \int_{\substack{r(z)=r(x)\\ s(z)<s(x)}} r(x)p(x)p(z)dz\,dx \overset{(29,30,31)}{=}$$

$$\frac{1}{2}\int_{\mathcal{X}} \int_{\substack{z\neq x\\ s(z)=s(x)}} \max\{r(x),r(z)\}p(x)p(z)dz\,dx + \frac{1}{2}\int_{\mathcal{X}}\int_{z=x} r(x)p(x)p(z)dz\,dx - \frac{1}{2}\int_{\mathcal{X}}\int_{r(z)=r(x)} r(x)p(x)p(z)dz\,dx\,.$$

Observe that

$$\min_{s:\mathcal{X}\to\mathbb{R}} F_1(s) = 0,$$

$$\min_{s:\mathcal{X}\to\mathbb{R}} F_2(s) = \frac{1}{2}\int_{\mathcal{X}}\int_{z=x} r(x)p(x)p(z)dz\,dx - \frac{1}{2}\int_{\mathcal{X}}\int_{r(z)=r(x)} r(x)p(x)p(z)dz\,dx\,,$$

and both minima are attained by a scoring function $s^*:\mathcal{X}\to\mathbb{R}$ if and only if conditions (26) and (27) hold for s^*. Also note that the conditions can be fulfilled, e.g. by any s^* such that

$$(\forall x,z\in\mathcal{X})\,(x\neq z \Rightarrow\; s^*(x)\neq s^*(z)\;\wedge\; r(x)<r(z)\Rightarrow s^*(x)<s^*(z))\,.$$

□

References

Stein, E. M. and Shakarchi, R. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press, 2009. ISBN 9781400835560. URL https://books.google.cz/books?id=2Sg3Vug65AsC.