Recommendations as Treatments: Supplementary Material

Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, Thorsten Joachims
Cornell University, Ithaca, NY, USA {TBS49, FA234, AS3354, NC475, TJ36}@CORNELL.EDU

Proof of Proposition 3.1

Proposition (Tail Bound for IPS Estimator)**.** *Let P be the independent Bernoulli probabilities of observing each entry. For any given $\hat{Y}$ and Y, with probability $1-\eta$, the IPS estimator $\hat{R}_{IPS}(\hat{Y}|P)$ does not deviate from the true $R(\hat{Y})$ by more than:*

$$\left|\hat{R}_{IPS}(\hat{Y}|P) - R(\hat{Y})\right| \leq \frac{1}{U\cdot I}\sqrt{\frac{\log\frac{2}{\eta}}{2}\sum_{u,i}\rho_{u,i}^2}\ ,$$

where $\rho_{u,i} = \frac{\delta_{u,i}(Y,\hat{Y})}{P_{u,i}}$ if $P_{u,i} < 1$, and $\rho_{u,i} = 0$ otherwise.

Proof. Hoeffding's inequality states that for independent bounded random variables $Z_1, ..., Z_n$ that take values in intervals of sizes $\rho_1, ..., \rho_n$ with probability 1 and for any $\epsilon > 0$

$$P\left(\left|\sum_k Z_k - E\left[\sum_k Z_k\right]\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-2\epsilon^2}{\sum_k \rho_k^2}\right)$$

Defining $P\left(Z_k = \frac{\delta_{u,i}(Y,\hat{Y})}{P_{u,i}}\right) = P_{u,i}$ and $P(Z_k = 0) = 1 - P_{u,i}$ relates Hoeffding's inequality to the IPS estimator and its expectation, which equals $R(\hat{Y})$ as shown earlier. This yields

$$P\left(\left|\hat{R}_{IPS}(\hat{Y}|P) - R(\hat{Y})\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-2\epsilon^2 U^2\cdot I^2}{\sum_{u,i}\rho_{u,i}^2}\right)$$

where $\rho_{u,i}$ is defined as in the statement of the proposition above. Solving for ϵ completes the proof. □

Proof of Theorem 4.2

Theorem (Propensity-Scored ERM Generalization Error Bound)**.** *For any finite hypothesis space of predictions $\mathcal{H} = \{\hat{Y}_1, ..., \hat{Y}_{|\mathcal{H}|}\}$ and loss $0 \leq \delta_{u,i}(Y,\hat{Y}) \leq \Delta$, the true risk $R(\hat{Y})$ of the empirical risk minimizer $\hat{Y}^{ERM}$ from H using the IPS estimator given training observations O from Y with independent Bernoulli propensities P is bounded with probability $1-\eta$ by*

$$R(\hat{Y}^{ERM}) \leq \hat{R}_{IPS}(\hat{Y}^{ERM}|P) + \frac{\Delta}{U\cdot I}\sqrt{\frac{\log(2|\mathcal{H}|/\eta)}{2}}\sqrt{\sum_{u,i}\frac{1}{P_{u,i}^2}} \quad (1)$$

Proof. Making a uniform convergence argument via Hoeffding and union bound yields:

$$\begin{aligned}
&P\left(\left|R(\hat{Y}^{ERM}) - \hat{R}_{IPS}(\hat{Y}^{ERM}|P)\right| \leq \epsilon\right) \geq 1-\eta \\
\Leftarrow\ &P\left(\max_{\hat{Y}_i}\left|R(\hat{Y}_i) - \hat{R}_{IPS}(\hat{Y}_i|P)\right| \leq \epsilon\right) \geq 1-\eta \\
\Leftrightarrow\ &P\left(\bigvee_{\hat{Y}_i}\left|R(\hat{Y}_i) - \hat{R}_{IPS}(\hat{Y}_i|P)\right| \geq \epsilon\right) < \eta \\
\Leftarrow\ &\sum_{i=1}^{|\mathcal{H}|} P\left(\left|R(\hat{Y}_i) - \hat{R}_{IPS}(\hat{Y}_i|P)\right| \geq \epsilon\right) < \eta \\
\Leftarrow\ &|\mathcal{H}|\cdot 2\exp\left(\frac{-2\epsilon^2}{\frac{\Delta^2}{U^2\cdot I^2}\sum_{u,i}\frac{1}{P_{u,i}^2}}\right) < \eta
\end{aligned}$$

Solving the last line for ϵ yields the desired result. □

Proof of Lemma 5.1

Lemma (Bias of IPS Estimator under Inaccurate Propensities)**.** *Let P be the marginal probabilities of observing an entry of the rating matrix Y, and let $\hat{P}$ be the estimated propensities such that $\hat{P}_{u,i} > 0$ for all u, i. The bias of the IPS estimator using $\hat{P}$ is*

$$\text{bias}\left(\hat{R}_{IPS}(\hat{Y}|\hat{P})\right) = \sum_{u,i}\frac{\delta_{u,i}(Y,\hat{Y})}{U\cdot I}\left[1 - \frac{P_{u,i}}{\hat{P}_{u,i}}\right]$$

Proof. Bias is defined as

$$\text{bias}\left(\hat{R}_{IPS}(\hat{Y}|\hat{P})\right) = R(\hat{Y}) - \mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}|\hat{P})\right],$$

where $R(\hat{Y})$ is the true risk of $\hat{Y}$ over the complete rating matrix. Expanding both terms yields

$$R(\hat{Y}) = \frac{1}{U\cdot I}\sum_{u,i}\delta_{u,i}(Y,\hat{Y}) \quad (2)$$

$$\mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}|\hat{P})\right] = \frac{1}{U\cdot I}\sum_{u,i}\frac{P_{u,i}}{\hat{P}_{u,i}}\delta_{u,i}(Y,\hat{Y}). \quad (3)$$

Rest follows after subtracting line (3) from (2). □

Proof of Theodem 5.2

Theorem (Propensity-Scored ERM Generalization Error Bound under Inaccurate Propensities)**.** *For any finite hypothesis space of predictions* $\mathcal{H} = \{\hat{Y}_1, ..., \hat{Y}_{|\mathcal{H}|}\}$*, the transductive prediction error of the empirical risk minimizer* $\hat{Y}^{ERM}$*, using the IPS estimator with estimated propensities* $\hat{P}$ *(*$\hat{P}_{u,i} > 0$*) and given training observations* O *from* Y *with independent Bernoulli propensities* P*, is bounded by:*

$$R(\hat{Y}^{ERM}) \leq \hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P}) + \frac{\Delta}{U \cdot I} \sum_{u,i} \left| 1 - \frac{P_{u,i}}{\hat{P}_{u,i}} \right| + \frac{\Delta}{U \cdot I} \sqrt{\frac{\log(2|\mathcal{H}|/\eta)}{2}} \sqrt{\sum_{u,i} \frac{1}{\hat{P}_{u,i}^2}} \; . \quad (4)$$

Proof. First, notice that we can write

$$\begin{aligned} R(\hat{Y}^{ERM}) &= R(\hat{Y}^{ERM}) - \mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P})\right] \\ &\quad + \mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P})\right] \\ &= \text{bias}\left(\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P})\right) \\ &\quad + \mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P})\right] \\ &\leq \frac{\Delta}{U \cdot I} \sum_{u,i} \left| 1 - \frac{P_{u,i}}{\hat{P}_{u,i}} \right| \\ &\quad + \mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P})\right] \end{aligned}$$

which follows from Lemma 5.1.

We are left to bound the following

$$\begin{aligned} & P\left(\left|\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P}) - \mathbb{E}_O\left[\hat{R}_{IPS}(\hat{Y}^{ERM}|\hat{P})\right]\right| \leq \epsilon\right) \\ & \geq 1 - \eta \\ & \Leftarrow |\mathcal{H}| \cdot 2 \exp\left(\frac{-2\epsilon^2}{\frac{\Delta^2}{U^2 \cdot I^2} \sum_{u,i} \frac{1}{\hat{P}_{u,i}^2}}\right) < \eta. \end{aligned}$$

The intermediate steps here are analogous to the steps in the proof of Theorem 4.2. Rearranging the terms and adding the bias gives the stated results. □

Propensity Estimation via Logistic Regression

In contrast to other discriminative models, logistic regression offers some attractive properties for propensity estimation.

Observation. *For the logistic propensity model, we observe that at optimality of the MLE estimate, the following two equations hold:*

$$\forall i : \sum_u O_{u,i} = \sum_u P_{u,i} \quad (5)$$

$$\forall u : \sum_i O_{u,i} = \sum_i P_{u,i}. \quad (6)$$

In other words, the logistic propensity model is able to learn well-calibrated marginal probabilities.

Proof. The log-likelihood function of the entire model after simplification is:

$$\begin{aligned} \ell(O|X, \phi) = & \sum_{(i,u):O_{u,i}=1} \left[w^T X_{u,i} + \beta_i + \gamma_u\right] \\ & - \sum_{i,u} \left[1 + e^{w^T X_{u,i} + \beta_i + \gamma_u}\right]. \end{aligned} \quad (7)$$

The gradient for bias term β_i (analogously for γ_u) for item i is given as

$$\frac{\partial \ell}{\partial \beta_i} = \sum_u O_{u,i} - \sum_u P_{u,i}. \quad (8)$$

Solving the gradient for zero yields the stated result. □