1　はじめに

本資料では、ランキング問題におけるオフ方策評価を扱う。

2　推薦枠内におけるランキングのオフ方策評価

2.1　KPI を設定する

KPI は、推薦枠において発生する総 CV 数とする。

2.2　データの観測構造をモデル化する

対象となる推薦枠においてコンバージョンが発生するまでの一連の流れを以下のように整理する。

1. ユーザーがサービスに訪問する。
2. 訪問ユーザーに対して、推薦枠内に複数のアイテムがランキング形式で表示される。
3. ユーザーの興味において、推薦枠内のいずれかのアイテムがクリックされる。クリックが発生しないケースもあり得る。
4. クリックされたアイテムについて、アイテム詳細ページに遷移した後、ユーザーの意思決定に基づきコンバージョンが発生するか否かが決定される。

ログデータ $\mathcal{D}$ は、以下のように定義する。

$$\mathcal{D} = \{(u_i, S_i, C_i, C_i R_i)\}_{i=1}^{n} \tag{2.1}$$

ここで、u_i はユーザー、S_i はランキング形式で表示されたアイテム、C_i は S_i の各アイテムがクリックされたかを表すベクトル、$C_i R_i$ は S_i の各アイテムが CV されたかを表すベクトルである。

ログデータ $\mathcal{D}$ の各要素は、独立同分布に従うとする。ログデータ $\mathcal{D}$ の従う確率分布は、

$$p(\mathcal{D}) = \prod_{i=1}^{n} p(u_i, S_i, C_i, C_i R_i) = \prod_{i=1}^{n} p(u_i) \pi_0(S_i | x_{u_i}) p(C_i, R_i | x_{u_i}, S_i) \tag{2.2}$$

とする。

2.3　解くべき問題を特定する

KPI は、推薦枠において発生する総 CV 数とした。これより、方策の価値は推薦枠における期待 CV 数を表す

$$V(\pi) = \mathbb{E}_{p(u)\pi(S|x_u)p(C,R|x_u,S)} \left[\sum_{k=1}^{K} C(k) R(k) \right] \tag{2.3}$$

とする。

2.4 観測データのみを用いて問題を解く方法を考える

既存の推薦モデル π_0 の下で収集されたログデータ D_0 のみを用いて、新たな推薦モデル π を導入した場合の方策の価値 $V(\pi)$ を推定するための推定量を考える。また、各推定量について、方策の価値 $V(\pi)$ に対するバイアスとバリアンスを導出し、推定量の性質を考察する。

2.4.1 IPS 推定量

IPS 推定量は、

$$\hat{V}_{\mathrm{IPS}}(\pi;\mathcal{D}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k) \tag{2.4}$$

で定義される推定量である。

まず、IPS 推定量のバイアスを求めるため、$p(\mathcal{D})$ の下での期待値を求める。

$$\begin{aligned}
\mathbb{E}_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{IPS}}(\pi;\mathcal{D})\right] &= \mathbb{E}_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k)\right] && (2.5)\\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{p(\mathcal{D})}\left[\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k)\right] && (2.6)\\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{p(u_i)\pi_0(S_i|x_{u_i})p(C_i,R_i|x_{u_i},S_i)}\left[\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k)\right] && (2.7)\\
&= \mathbb{E}_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\sum_{k=1}^{K}C(k)R(k)\right] && (2.8)\\
&= \mathbb{E}_{p(u)\pi_0(S|x_u)}\left[\mathbb{E}_{p(C,R|x_u,S)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\sum_{k=1}^{K}C(k)R(k)\right]\right] && (2.9)\\
&= \mathbb{E}_{p(u)\pi_0(S|x_u)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right] && (2.10)\\
&= \mathbb{E}_{p(u)}\left[\mathbb{E}_{\pi_0(S|x_u)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right]\right] && (2.11)\\
&= \mathbb{E}_{p(u)}\left[\sum_{S\in\mathcal{S}}\pi_0(S|x_u)\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right] && (2.12)\\
&= \mathbb{E}_{p(u)}\left[\sum_{S\in\mathcal{S}}\pi(S|x_u)\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right] && (2.13)\\
&= \mathbb{E}_{p(u)}\left[\mathbb{E}_{\pi(S|x_u)}\left[\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right]\right] && (2.14)\\
&= \mathbb{E}_{p(u)\pi(S|x_u)p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right] && (2.15)
\end{aligned}$$

$$= V(\pi) \tag{2.16}$$

これより、IPS 推定量のバイアスは 0 である。

次に、IPS 推定量のバリアンスを求める。

$$\begin{aligned}
\mathbb{V}_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{IPS}}(\pi;\mathcal{D})\right] &= \mathbb{V}_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k)\right] && (2.17)\\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}_{p(\mathcal{D})}\left[\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k)\right] && (2.18)\\
&= \frac{1}{n}\mathbb{V}_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\sum_{k=1}^{K}C(k)R(k)\right] && (2.19)\\
&= \frac{1}{n}\Bigg(\mathbb{E}_{p(u)\pi_0(S|x_u)}\left[\mathbb{V}_{p(C,R|x_u,S)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\sum_{k=1}^{K}C(k)R(k)\right]\right] && (2.20)\\
&\qquad + \mathbb{V}_{p(u)\pi_0(S|x_u)}\left[\mathbb{E}_{p(C,R|x_u,S)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\sum_{k=1}^{K}C(k)R(k)\right]\right]\Bigg)\\
&= \frac{1}{n}\Bigg(\mathbb{E}_{p(u)\pi_0(S|x_u)}\left[\left(\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\right)^2\mathbb{V}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right] && (2.21)\\
&\qquad + \mathbb{V}_{p(u)\pi_0(S|x_u)}\left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)}\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right]\right]\Bigg)
\end{aligned}$$

アイテム数が多い場合にはランキングの組み合わせも多くなるため、$\pi_0(S|x_u)$ が小さくなることでバリアンスが大きくなることが予想される。

2.4.2 SNIPS 推定量

SNIPS 推定量は、

$$\hat{V}_{\mathrm{SNIPS}}(\pi;\mathcal{D}) = \frac{1}{\sum_{i=1}^{n}\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}}\sum_{i=1}^{n}\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k) \tag{2.22}$$

で定義される推定量である。

IPS 推定量と異なるためバイアスは 0 にならないが、重要度重みが正規化されることでバリアンスを低減できることが期待される。

2.4.3 Clipped IPS 推定量

Clipped IPS 推定量は、

$$\hat{V}_{\mathrm{cIPS}}(\pi;\mathcal{D},M) = \frac{1}{n}\sum_{i=1}^{n}\min\left\{\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}, M\right\}\sum_{k=1}^{K}C_i(k)R_i(k) \tag{2.23}$$

で定義される推定量である。

まず、Clipped IPS 推定量のバイアスは、

$$\begin{aligned}\mathrm{Bias}\left[\hat{V}_{\mathrm{cIPS}}(\pi;\mathcal{D},M)\right] &= \mathbb{E}_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{cIPS}}(\pi;\mathcal{D},M)\right] - V(\pi) &(2.24)\\ &= \mathbb{E}_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\min\left\{\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})},M\right\}\sum_{k=1}^{K}C_i(k)R_i(k)\right] &(2.25)\\ &\quad -\mathbb{E}_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\sum_{k=1}^{K}C_i(k)R_i(k)\right]\\ &= \mathbb{E}_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\min\left\{\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})},M\right\}-\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\right)\sum_{k=1}^{K}C_i(k)R_i(k)\right] &(2.26)\end{aligned}$$

となる。ここでは、$V(\pi)=\mathbb{E}_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{IPS}}(\pi;\mathcal{D})\right]$ であることを用いた。

$$\left(\min\left\{\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})},M\right\}-\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})}\right)\leq 0 \tag{2.27}$$

であるから、Clipped IPS 推定量は不偏推定量ではなく、バイアスは 0 以下の値をとることがわかる。

次に、Clipped IPS 推定量のバリアンスは、IPS 推定量と同様にして

$$\begin{aligned}V_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{cIPS}}(\pi;\mathcal{D},M)\right] &= V_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\min\left\{\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})},M\right\}\sum_{k=1}^{K}C_i(k)R_i(k)\right] &(2.28)\\ &= \frac{1}{n^2}\sum_{i=1}^{n}V_{p(\mathcal{D})}\left[\min\left\{\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})},M\right\}\sum_{k=1}^{K}C_i(k)R_i(k)\right] &(2.29)\\ &= \frac{1}{n}V_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)}\left[\min\left\{\frac{\pi(S|x_u)}{\pi_0(S|x_u)},M\right\}\sum_{k=1}^{K}C(k)R(k)\right] &(2.30)\end{aligned}$$

となる。M で抑えることで、IPS 推定量よりもバリアンスが小さくなることがわかる。

2.4.4 IIPS 推定量

IIPS 推定量は、

$$\hat{V}_{\mathrm{IIPS}}(\pi;\mathcal{D})=\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)}C_i(k)R_i(k) \tag{2.31}$$

で定義される推定量である。

まず、IIPS 推定量のバイアスを求めるため、$p(\mathcal{D})$ の下での期待値を求める。

$$\begin{aligned}\mathbb{E}_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{IIPS}}(\pi;\mathcal{D})\right] &= \mathbb{E}_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)}C_i(k)R_i(k)\right] &(2.32)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)}\left[\sum_{k=1}^{K}\frac{\pi(S(k)|x_u,k)}{\pi_0(S(k)|x_u,k)}C(k)R(k)\right] &(2.33)\end{aligned}$$

$$= \mathbb{E}_{p(u)}\left[\mathbb{E}_{\pi_0(S|x_u)}\left[\mathbb{E}_{p(C,R|x_u,S)}\left[\sum_{k=1}^{K}\frac{\pi(S(k)|x_u,k)}{\pi_0(S(k)|x_u,k)}C(k)R(k)\right]\right]\right] \tag{2.34}$$

$$= \mathbb{E}_{p(u)}\left[\mathbb{E}_{\pi_0(S|x_u)}\left[\sum_{k=1}^{K}\frac{\pi(S(k)|x_u,k)}{\pi_0(S(k)|x_u,k)}\mathbb{E}_{p(C,R|x_u,S)}\left[C(k)R(k)\right]\right]\right] \tag{2.35}$$

$$= \mathbb{E}_{p(u)}\left[\mathbb{E}_{\pi_0(S|x_u)}\left[\sum_{k=1}^{K}\frac{\pi(S(k)|x_u,k)}{\pi_0(S(k)|x_u,k)}\mathbb{E}_{p(C,R|x_u,S(k))}\left[C(k)R(k)\right]\right]\right] \tag{2.36}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{S\in\mathcal{S}}\pi_0(S|x_u)\sum_{k=1}^{K}\frac{\pi(S(k)|x_u,k)}{\pi_0(S(k)|x_u,k)}\mathbb{E}_{p(C,R|x_u,S(k))}\left[C(k)R(k)\right]\right] \tag{2.37}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{S\in\mathcal{S}}\pi_0(S|x_u)\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\frac{\pi(a|x_u,k)}{\pi_0(a|x_u,k)}\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\mathbb{1}\{S(k)=a\}\right] \tag{2.38}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\frac{\pi(a|x_u,k)}{\pi_0(a|x_u,k)}\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\sum_{S\in\mathcal{S}}\pi_0(S|x_u)\mathbb{1}\{S(k)=a\}\right] \tag{2.39}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\frac{\pi(a|x_u,k)}{\pi_0(a|x_u,k)}\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\pi_0(a|x_u,k)\right] \tag{2.40}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\pi(a|x_u,k)\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\right] \tag{2.41}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\left(\sum_{S\in\mathcal{S}}\pi(S|x_u)\mathbb{1}\{S(k)=a\}\right)\right] \tag{2.42}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\sum_{S\in\mathcal{S}}\pi(S|x_u)\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\mathbb{1}\{S(k)=a\}\right] \tag{2.43}$$

$$= \mathbb{E}_{p(u)}\left[\sum_{S\in\mathcal{S}}\pi(S|x_u)\sum_{k=1}^{K}\sum_{a\in\mathcal{A}}\mathbb{E}_{p(C,R|x_u,a)}\left[C(k)R(k)\right]\mathbb{1}\{S(k)=a\}\right] \tag{2.44}$$

$$= \mathbb{E}_{p(u)\pi(S|x_u)}\left[\sum_{k=1}^{K}\mathbb{E}_{p(C,R|x_u,S(k))}\left[C(k)R(k)\right]\right] \tag{2.45}$$

$$= \mathbb{E}_{p(u)\pi(S|x_u)}\left[\sum_{k=1}^{K}\mathbb{E}_{p(C,R|x_u,S)}\left[C(k)R(k)\right]\right] \tag{2.46}$$

$$= \mathbb{E}_{p(u)\pi(S|x_u)p(C,R|x_u,S)}\left[\sum_{k=1}^{K}C(k)R(k)\right] \tag{2.47}$$

$$= V(\pi) \tag{2.48}$$

これより、IIPS 推定量のバイアスは、独立性が満たされる状況では 0 である。

次に、IIPS 推定量のバリアンスを求める。

$$V_{p(\mathcal{D})}\left[\hat{V}_{\mathrm{IIPS}}(\pi;\mathcal{D})\right] = V_{p(\mathcal{D})}\left[\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)}C_i(k)R_i(k)\right] \tag{2.49}$$

$$= \frac{1}{n} V_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)} \left[\sum_{k=1}^{K} \frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)} C_i(k) R_i(k) \right] \tag{2.50}$$

$$= \frac{1}{n} \Bigg(\mathbb{E}_{p(u)\pi_0(S|x_u)} \left[V_{p(C,R|x_u,S)} \left[\sum_{k=1}^{K} \frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)} C_i(k) R_i(k) \right] \right] \tag{2.51}$$

$$+ V_{p(u)\pi_0(S|x_u)} \left[\mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^{K} \frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)} C_i(k) R_i(k) \right] \right] \Bigg) \tag{2.52}$$

IPS 推定量のような方策のランキングに対する確率ではなく、アイテムに対する確率になっていることから、IPS 推定量よりもバリアンスが小さくなることを期待できる。

2.4.5 SNIIPS 推定量

SNIIPS 推定量は、

$$\hat{V}_{\mathrm{SNIIPS}}(\pi; \mathcal{D}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \frac{\frac{\pi(S_i(k)|x_{u_i},k)}{\pi_0(S_i(k)|x_{u_i},k)}}{\sum_{j=1}^{n} \frac{\pi(S_j(k)|x_{u_j},k)}{\pi_0(S_j(k)|x_{u_j},k)}} C_i(k) R_i(k) \tag{2.53}$$

で定義される推定量である。

IIPS 推定量と異なるためバイアスは 0 にならないが、重要度重みが正規化されることでバリアンスを低減できることが期待される。

2.4.6 Naive 推定量

Naive 推定量は、

$$\hat{V}_{\mathrm{Naive}}(\pi; \mathcal{D}) = \frac{1}{\sum_{i=1}^{n} \sum_{k=1}^{K} \pi(S_i(k)|x_{u_i},k)} \sum_{i=1}^{n} \sum_{k=1}^{K} \pi(S_i(k)|x_{u_i},k) C_i(k) R_i(k) \tag{2.54}$$

で定義される推定量である。

これは、π が確定的な方策の場合を考えると、ログを収集した時の方策 π_0 が出したランキングの k 番目のアイテムと、新たな方策 π が出したランキングの k 番目のアイテムが一致した時のみ、その報酬を加算して平均した値となる。実際、i 番目のサンプルに対して π がランキング T_i を確定的に出すとすると、

$$\hat{V}_{\mathrm{Naive}}(\pi; \mathcal{D}) = \frac{1}{\sum_{i=1}^{n} \sum_{k=1}^{K} \pi(S_i(k)|x_{u_i},k)} \sum_{i=1}^{n} \sum_{k=1}^{K} \pi(S_i(k)|x_{u_i},k) C_i(k) R_i(k) \tag{2.55}$$

$$= \frac{1}{\sum_{i=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{S_i(k) = T_i(k)\}} \sum_{i=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{S_i(k) = T_i(k)\} C_i(k) R_i(k) \tag{2.56}$$

バイアスはあるが、バリアンスを低減する効果があるようである。

2.4.7　DM 推定量

DM 推定量は、

$$\hat{V}_{\mathrm{DM}}(\pi;\mathcal{D}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\mathbb{E}_{\pi(S|x_{u_i})}\left[\hat{q}(x_{u_i}, e_{S(k)})\right] \tag{2.57}$$

で定義される推定量である。ここで、

$$q(x_{u_i}, e_{S(k)}) = \mathbb{E}\left[C(k)R(k)|x_u, e_{S(k)}\right] \tag{2.58}$$

である。すなわち、$q(x_{u_i}, e_{S(k)})$ はユーザー u がアイテム $S(k)$ を CV する確率である。$\hat{q}$ は、ログデータから機械学習モデルを学習させることで得る。

DM 推定量は、$\hat{q}$ の予測精度に依存したバイアスを持つが、バリアンスが小さい利点がある。

2.4.8　DR 推定量

3　検索におけるランキングのオフ方策評価

3.1　KPI を設定する

KPI は、推薦枠において発生する総 CV 数とする。

3.2　データの観測構造をモデル化する

対象となる推薦枠においてコンバージョンが発生するまでの一連の流れを以下のように整理する。

1. ユーザーがサービスに訪問する。
2. ユーザーに検索条件を提示する。
3. ユーザーが検索条件を選択する。
4. 訪問ユーザーに対して、推薦枠内に複数のアイテムがランキング形式で表示される。
5. ユーザーの興味において、推薦枠内のいずれかのアイテムがクリックされる。クリックが発生しないケースもあり得る。
6. クリックされたアイテムについて、アイテム詳細ページに遷移した後、ユーザーの意思決定に基づきコンバージョンが発生するか否かが決定される。

ログデータ $\mathcal{D}$ は、以下のように定義する。

$$\mathcal{D} = \{(u_i, q_i, l_i, S_i, C_i, C_i R_i)\}_{i=1}^{n} \tag{3.1}$$

ログデータ $\mathcal{D}$ の各要素は、独立同分布に従うとする。ログデータ $\mathcal{D}$ の従う確率分布は、

$$p(\mathcal{D}) = \prod_{i=1}^{n} p(u_i, q_i, l_i, S_i, C_i, C_i R_i) = \prod_{i=1}^{n} p(u_i)\pi_0^{q}(q_i|x_{u_i})p(l_i|x_{u_i}, q_i)\pi_0^{s}(S_i|x_{u_i}, q_i, l_i)p(C_i, R_i|x_{u_i}, q_i, l_i, S_i) \tag{3.2}$$

とする。

3.3　解くべき問題を特定する

KPI は、推薦枠において発生する総 CV 数とした。これより、方策の価値は推薦枠における期待 CV 数を表す

$$V(\pi) = \mathbb{E}_{p(u)\pi(S|x_u)p(C,R|x_u,S)}\left[\sum_{k=1}^{K} C(k)R(k)\right] \tag{3.3}$$

とする。

4　プラットフォーム全体におけるランキングのオフ方策評価