

問題

あるポータルサイトに r 個のページが掲載されています。各ページ i ($i = 1, \dots, r$, $r \geq 2$) の一日当たりのアクセス数 X_i はパラメータ λ_i のポアソン分布 $\text{Po}(\lambda_i)$ に互いに独立に従うとします。 $X \sim \text{Po}(\lambda)$ の確率関数は

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

です。以下の各問に答えてください。

[1-1] パラメータ λ のポアソン分布の平均及び分散の値として正しいものを以下の選択肢から選び、それぞれ `submit.yml` のキー `1-1-mean`, `1-1-variance` に記入してください。

1. λ
2. 2λ
3. λ^2
4. $\sqrt{\lambda}$

いま、ある一日の各ページのアクセス数 $X = (X_1, \dots, X_r)^\top$ を元に平均アクセス数の同時推定を考えます。 $\delta(X) : \mathbb{N}^r \rightarrow \mathbb{R}_{>0}^r$ をその推定量、 $\delta_i(X)$ を $\delta(X)$ の i 番目の次元、つまりページ i に対するアクセス数の推定量とします。推定量 δ の良さを以下のリスク関数 $R(\lambda, \delta)$ で評価することにします：

$$R(\lambda, \delta) = \mathbb{E}_X [L(\lambda, \delta(X))], \text{ where } L(\lambda, \delta(X)) = \sum_{i=1}^r \frac{(\lambda_i - \delta_i(X))^2}{\lambda_i}.$$

ただし $\mathbb{E}_X [f(X)]$ は確率変数 X に関する $f(x)$ の期待値を表します。

[1-2] 最尤推定量 $\delta_{\text{ML}}(X) = X = (X_1, \dots, X_r)^\top$ に対するリスク関数 $R(\lambda, \delta_{\text{ML}})$ の値として正しいものを以下の選択肢から選び、`submit.yml` のキー `1-2` に記入してください。ただし、 $\|\lambda\|^2 = \sum_{i=1}^r \lambda_i^2$ と定義します。

1. $\|\lambda\|$
2. $\|\lambda\|^2$
3. $\sqrt{\|\lambda\|}$
4. 1
5. $1/2$
6. r
7. r^2
8. $r/\|\lambda\|$

次に最尤推定量の代わりに以下の推定量 δ_H の挙動を調べます：

$$\delta_H(X) = (1 - H(Z)) X$$

ただし $Z = \sum_{i=1}^r X_i$ とし、 $H : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ は以下の2つの条件を満たす関数とします：

- 任意の $x \in \mathbb{R}_{>0}$ に対して $0 \leq H(x) < 2 \frac{r-1}{x+r-1}$

- $xH(x)$ は単調非減少関数

[2] Z が与えられた下での X_i の条件つき期待値及び分散の値として正しいものを以下の選択肢から選び、それぞれ `submit.yml` のキー `2-mean`, `2-variance` に記入してください。ただし $\Lambda = \sum_{i=1}^r \lambda_i$ とします。

1. Z
2. $Z\lambda_i$
3. $Z\lambda_i/\Lambda$
4. $Z(\lambda_i/\Lambda)(1 - \lambda_i/\Lambda)$
5. λ_i/Λ
6. $(\lambda_i/\Lambda)(1 - \lambda_i/\Lambda)$

[3-1] 以下の (a), (b), (c) に当てはまる数式として正しいものを以下の選択肢から選び、それぞれ `submit.yml` のキー `3-1-a`, `3-1-b`, `3-1-c` に記入してください。

$$R(\lambda, \delta_H) = (a) + \mathbb{E}_Z[(b) \cdot H(Z) + (c) \cdot H^2(Z)]$$

1. r
2. r^2
3. Z
4. Z/Λ
5. $\Lambda - Z - r + 1$
6. $(\Lambda - Z - r + 1)/\Lambda$
7. $2Z(\Lambda - Z - r + 1)/\Lambda$
8. $\Lambda(Z + r - 1)$
9. $Z(Z + r - 1)/\Lambda$

[3-2] これまでの結果を用いて $R(\lambda, \delta_H) \leq R(\lambda, \delta_{ML})$ を示します。以下の証明内の空欄 (d) ~ (g) に入る適切な数式を選択肢から選び、キー `3-2-d`, `3-2-e`, `3-2-f`, `3-2-g` に記入して下さい。各選択肢は複数回使われる可能性もあります。また、 $f(x)$ を単調非減少関数、 $g(x)$ を単調非増加関数とすると

$$\mathbb{E}_X[f(X)g(X)] \leq \mathbb{E}_X[f(X)]\mathbb{E}_X[g(X)]$$

が成り立ちます。

(証明) [3-1] の結果と H の条件から

$$R(\lambda, \delta_H) \leq (a) + 2\mathbb{E}_Z\left[\frac{(d)(e)}{(f)}\{(f) - (d)\}\right]$$

が成り立ちます。第二項の期待値の項について、与えられた不等式を用いることで

$$\mathbb{E}_Z\left[\frac{(d)(e)}{(f)}\{(f) - (d)\}\right] \leq \mathbb{E}_Z\left[\frac{(d)(e)}{(f)}\right]\mathbb{E}_Z[(f) - (d)]$$

が得られます。ここで Z が従う分布を考えることで、 $\mathbb{E}_Z[(f) - (d)]$ は (g) になることがわかります。

以上より、 $R(\lambda, \delta_H) \leq R(\lambda, \delta_{ML})$ であることが示されました。

1. Λ
2. $\|\lambda\|$
3. Z
4. H
5. r
6. 0
7. 1