

順序統計量に対する事後予測分布の優越性評価

東京理科大学 西 航平

東京理科大学 黒沢 健

本稿では、順序統計量に関する未観測データのベイズ型予測分布として、plug-in 型予測分布 f_δ と事後予測分布 p を与え、それらの優越性について議論する。一般に、plug-in 型予測分布よりも事後予測分布を予測分布として使用することが推奨されている [1] が、必ずしも優越的であるわけではない。そこで順序統計量の未観測データの予測分布 f_δ, p の優越性について検討する。

真の分布 f と評価したい分布 p の間の損失を Kullback-Leibler 損失 (KL 損失) で測る。 f と p の KL 損失は

$$L_{KL}(f, p) = \int_Y f(y|\theta) \log \frac{f(y|\theta)}{p(y|\theta)} dy$$

で与えられるが、KL 損失は観測データ $\mathbf{t}$ と θ に依存するため優越性を評価することが難しい。そこで期待損失 (リスク) を求めることで評価を行う。リスクは観測データ $\mathbf{t}$ の分布 $f(\mathbf{t}|\theta)$ 上での KL 損失の期待値として以下のように定義される。これは未知パラメータ θ の関数であるため、 θ に関して一様に評価できることが望まれる。

$$R(\theta, p) = \int_T L_{KL}(f, p) f(\mathbf{t}|\theta) d\mathbf{t}.$$

plug-in 型予測分布に関するリスク $R(\theta, \delta)$ と事後予測分布に関するリスク $R(\theta, p)$ の差により優越性を評価する。

本検討では $X_1, X_2, \dots, X_n \stackrel{i.i.d.}{\sim} \text{Exp}(\lambda)$ とし、 $X_1, X_2, \dots, X_n$ の順序統計量 $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ ($X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$) の内の $\mathbf{T}_{n-1} = (X_{(1)}, \dots, X_{(n-1)})$ が $\mathbf{t}_{n-1} = (t_1, \dots, t_{n-1})$ と観測されたもとで、未観測値 $Y = X_{(n)} | (\mathbf{T}_{n-1} = \mathbf{t}_{n-1})$ を予測する予測分布を考える。このときの Y の真のモデルは

$$f_Y(y|\mathbf{t}_{n-1}) = \lambda \exp(-\lambda(y - t_{n-1})) \quad (y \geq t_{n-1})$$

で与えられる。ここで λ の事前分布 $\pi(\lambda)$ をガンマ分布 $Ga(\alpha, \beta)$ として、plug-in 型予測分布と事後予測分布を比較する。ただし plug-in する λ の推定値は最尤推定値 $\delta(\mathbf{t}_{n-1})$ を考える。

- 推定値 δ による plug-in 型予測分布

$$f_\delta := f(y|\delta, \mathbf{t}_{n-1}; \alpha, \beta) = \delta \exp(\delta(y - t_{n-1})).$$

- 事後予測分布

$$p_Y := p_Y(y|\mathbf{t}_{n-1}; \alpha, \beta) = (\alpha + n - 1) \frac{(\beta + t_1 + \dots + t_{n-1})^{(\alpha+n-1)}}{(y + \beta + t_1 + \dots + t_{n-1})^{\alpha+n}}.$$

リスクの差は以下のように求まる。

$$R(\lambda, f_\delta) - R(\lambda, p) = \int_0^\infty \dots \int_{t_{n-2}}^\infty (L_{KL}(f, f_\delta) - L_{KL}(f, p)) dt_{n-1} \dots dt_1.$$

特に最尤推定値 δ を plug-in する場合は変数変換を施すことでリスクの差は以下のように求まることを示した。

$$R(\theta, f_\delta) - R(\theta, p) = E^{Z_1} E^{Z_2} \dots E^{Z_{n-1}} [G_1(Z_1 + \dots + Z_{n-1})] \quad (Z_1, \dots, Z_{n-1} \stackrel{i.i.d.}{\sim} \text{Exp}(\lambda)).$$

ただし

$$G_1(z) = \log \frac{\alpha + n - 1}{n - 1} - \log(\beta + z) + \log z + (\alpha + n - 1 + 1) e^{\lambda(\beta+z)} \text{Ei}(-\lambda(\beta+z)) + \frac{n-1}{z\lambda}.$$

また、 $(\alpha, \beta) \rightarrow (0, 0)$ とした場合において明示的にリスクの差を与え、事後予測分布が一様に優位性を持つことを示した。本研究は JSPS 科研費 JP18K11200 の助成を受けたものである。

参考文献

- [1] F. Komaki. On asymptotic properties of predictive distributions. *Biometrika*, Vol. 83, pp. 299–313, 1996.