問 1

未知の母数の推定量として標本平均、中央値、トリム平均、加重平均、幾何平均について簡単に説明せよ。なお、下の観測値 X の標本平均、中央値とトリム平均（小さい方と大きい方の 1 個を捨てる）を求めよ。

観測値 X：

1.651	3.138	1.243	1.732	2.432	0.721

解答

- 標本平均：

$$\hat{\theta}=\frac{X_1+X_2+...+X_n}{n}$$

観察される確率変数の算術平均で定義される推定量である。

- 中央値：

$$\hat{\theta}=\left(X_1,X_2,...,X_n\text{を大きさの順に並べたとき真ん中に来る値}\right)$$

$$=\begin{cases} X_{[(n+1)/2]} & ,n\text{ が奇数の場合} \\ \dfrac{X_{[n/2]}+X_{[n/2+1]}}{2} & ,n\text{ が偶数の場合} \end{cases}$$

ただし、$\left\{X_{[i]}\right\}$ は観測値 $\left\{X_i\right\}$ を小さい順に並べ換えたものである。

- トリム平均：

$$\hat{\theta}=(X_1,X_2,...,X_n\text{ を小さい順に並べ、小さいほうの }m\text{ 個と大きいほうの }m\text{ 個を捨てた算術平均})$$

$$=\frac{1}{n-2m}\sum_{i=m+1}^{n-m}X_i$$

極端に大きな、あるいは小さな値として現れる観測値を捨てることによって、そうした特異的な値による誤差の影響を取り除いたうえで平均を推定しようというものである。

- 加重平均：

$$\hat{\theta}=\sum_{i=1}^{n}c_iX_i\text{ , ただし, }\sum_{i=1}^{n}c_i=1\text{ とする。}$$

各観測の重みを変えて平均を取ったものであり、観測値ごとに重要性が異なると考えていることになる。

- 幾何平均：

$\hat{\theta}=\left(X_1X_2...X_n\right)^{1/n}$　ただし $X_1,X_2,...,X_n$ を正の値とする。一般に等比数列の真ん中の値を求めるための方法である。

- 観測値 X の標本平均、中央値とトリム平均はそれぞれ 1.820、1.692 と 1.765 である。

以上

問 2

$\hat{\theta}_1,\ \hat{\theta}_2,\ ...,\ \hat{\theta}_k$ が、いずれも不偏推定量であり、その分散が等しく σ^2 である。すなわち

$$V(\hat{\theta}_1)=V(\hat{\theta}_2)=...=V(\hat{\theta}_k)=\sigma^2$$

とする。このとき、これらの推定量の算術平均を

$$\hat{\theta}^*=\frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i$$

とおけば

$$E(\hat{\theta}^*)=\theta\quad\text{（不偏性は保たれる）}$$

$$V(\hat{\theta}^*)\le\sigma^2\quad\text{（分散が小さくなる）}$$

が成立することを証明せよ。

解答

平均値の線形性から不偏性は明らかである。

$$E(\hat{\theta}^*)=E(\frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i)=\frac{1}{k}\sum_{i=1}^{k}E(\hat{\theta}_i)=\frac{1}{k}\sum_{i=1}^{k}\theta=\theta$$

分散については：

$$\sum_{i=1}^{k}\left(\hat{\theta}_i-\theta\right)^2=\sum_{i=1}^{k}\left(\hat{\theta}_i-\hat{\theta}^*+\hat{\theta}^*-\theta\right)^2$$

$$=\sum_{i=1}^{k}\left(\hat{\theta}_i-\hat{\theta}^*\right)^2+2\left(\hat{\theta}^*-\theta\right)\sum_{i=1}^{k}\left(\hat{\theta}_i-\hat{\theta}^*\right)+\sum_{i=1}^{k}\left(\hat{\theta}^*-\theta\right)^2$$

$$\left(\hat{\theta}^*=\frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i\ \text{ より }\ \sum_{i=1}^{k}\left(\hat{\theta}_i-\hat{\theta}^*\right)=0\right)$$

$$=\sum_{i=1}^{k}\left(\hat{\theta}_i-\hat{\theta}^*\right)^2+k\left(\hat{\theta}^*-\theta\right)^2$$

が成り立つことから、両方の平均をとって

$$\sum_{i=1}^{k}E\left[\left(\hat{\theta}_i-\theta\right)^2\right]=\sum_{i=1}^{k}E\left[\left(\hat{\theta}_i-\hat{\theta}^*\right)^2\right]+kE\left[\left(\hat{\theta}^*-\theta\right)^2\right]$$

$$\sum_{i=1}^{k}V\left(\hat{\theta}_i\right)=\sum_{i=1}^{k}E\left[\left(\hat{\theta}_i-\hat{\theta}^*\right)^2\right]+kV\left(\hat{\theta}^*\right)$$

$$k\sigma^2=\sum_{i=1}^{k}E\left[\left(\hat{\theta}_i-\hat{\theta}^*\right)^2\right]+kV\left(\hat{\theta}^*\right)\ge kV\left(\hat{\theta}^*\right)$$

となり、算術平均の分散のほうが小さくなることがわかる。

以上