The Chinese University of Hong Kong, Shenzhen
DDA4002 2024-2025 Spring

Solutions to Assignment 3

Yinqi (Marcus) Zhong
121090847

March 19, 2025

1 Theoretic Questions

1.1 Input Analysis

#1 **(Input Analysis)** Consider the shifted (two-parameter) exponential distribution, which has density function

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-(x-\gamma)/\beta} & \text{if } x \geq \gamma \\ 0 & \text{otherwise} \end{cases},$$

for $\beta > 0$ and any real number γ. Given a sample $X_1, X_2, \ldots, X_n$ of i.i.d. random values from this distribution, find formulas for the joint MLEs $\hat{\gamma}$ and $\hat{\beta}$. *Hint: Remember that γ cannot exceed any X_i.*

Figure 1: Question Setting

First, we need to derive the likelihood function of this distribution.

$$L(X, \gamma, \beta) = \prod_{i=1}^{n} \beta \cdot exp(-(x_i - \gamma)\beta)$$

The log-likelihood would be:

$$L(X, \gamma, \beta) = \log[\prod_{i=1}^{n} \frac{1}{\beta} \cdot exp(-(x_i - \gamma)\beta)] = -nlog(\beta) - \sum_{i=1}^{n} \frac{(x_i - \gamma)}{\beta}$$

Given hint that $\gamma \leq \min\{x_1, ..., x_n\}$. The MLE solution is equal to the constrained linear optimization for X, γ, β:

$$\begin{aligned} \max_{\gamma,\beta} \quad & l(X, \gamma, \beta) = -nlog(\beta) - \sum_{i=1}^{n} \frac{(x_i - \gamma)}{\beta} \\ \text{s.t.} \quad & \gamma \leq \min\{x_1, \ldots, x_n\} \end{aligned}$$

Take derivative with respect to γ and β:

$$\frac{\partial l}{\partial \gamma} = \frac{n}{\beta} > 0$$

$$\frac{\partial l}{\partial \beta} = -\frac{n}{\beta} + \frac{\sum_{i=1}^{n} x_i - n\gamma}{\beta^2}$$

Therefore, l is monotonically increasing at the realm of γ, so the MLE of γ is achieved at the maximum value of γ. So,

$$\hat{\gamma} = \min\{x_1, ...x_n\}$$

Applying the formula of $\hat{\gamma}$ and set the equation to be 0, we got

$$\hat{\beta} = \bar{x} - \hat{\gamma} = \bar{x} - \min\{x_1, ...x_n\}$$

1.2 Output Analysis

#2 **(Output Analysis)** Let p be the probability of success of some random experiment. We want to estimate the mean number of experiments to obtain the first success, i.e. $1/p$. Suppose we have n i.i.d. outcomes $X_1, ..., X_n$ of the random experiment via simulation. In detail, $X_i = 1$(the ith experiment succeeds). Derive an estimate for $1/p$ from the simulation data and a 95% confidence interval using Delta method.

Figure 2: Question Setting

First, we need to derive the MLE for p, which is:

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \hat{x}_n$$

So, the estimate for $\frac{1}{p}$ is $\frac{1}{\hat{x}_n}$.

The variance of the estimator should be calculated using Central Limit Theorem. Where we know:

$$\sqrt{n}(\hat{x}_n - \bar{x}) \xrightarrow{d} N(0, \sigma^2), \quad \text{where} \quad \sigma 2 = Var(X_1) = p(1-p).$$

The MLE for σ is $\hat{\sigma} = \hat{p}(1-\hat{p}) = \hat{x}(1-\hat{x})$.

To implement the Delta Method, we may choose the smooth function $h(.) = \frac{1}{p}$, so $h'(p) = -\frac{1}{p^2}$.

So, for the **Delta Method**, we have

$$\sqrt{n}(\frac{1}{\hat{p}} - \frac{1}{p}) \equiv -\frac{1}{p^2}\sqrt{n} \cdot (\hat{p} - p) \xrightarrow{d} N(0, \frac{p(1-p)}{p^4})$$

So, using this distribution, we can derive the 95% CI for $\frac{1}{p}$ as

$$\left[\frac{1}{\hat{x}_n} - \phi_{0.975} \frac{\sqrt{\hat{x}_n(1-\hat{x}_n)}}{\hat{x}_n^2 \sqrt{n}}, \frac{1}{\hat{x}_n} + \phi_{0.975} \frac{\sqrt{\hat{x}_n(1-\hat{x}_n)}}{\hat{x}_n^2 \sqrt{n}} \right]$$

2 Numerical Experiments

2.1 Using True Value of Variance

To generate the 95% CI using the true value of limiting variance, we can follow these steps:

1. We generate $Z_1, \ldots, Z_n$ for $n = 500$ with $a = 0.5$.
2. We compute $\bar{Z}_n = \frac{1}{n}\sum_{t=1}^{n} Z_t$.

3. We construct the 95% CI

$$\left[\overline{Z}_n - 1.96\sqrt{\frac{\sigma^2}{n}}, \overline{Z}_n + 1.96\sqrt{\frac{\sigma^2}{n}}\right], \text{ where } \sigma^2 = \frac{1}{1-a^2} \cdot \frac{1+a}{1-a}.$$

4. We check whether 0 lies in that interval.
5. Repeat above steps for 1000 rounds; then report the empirical coverage, i.e., the proportion of times 0 is covered.

The empirical coverage probability (Method 1) = 0.956.

2.2 Using Sample Variance

To use sample variance, we need to adjust the above step 3. The general algorithm would be as follows.

1. We generate $Z_1, \ldots, Z_n$ for $n = 500$ with $a = 0.5$.
2. We compute $\bar{Z}_n = \frac{1}{n}\sum_{t=1}^{n} Z_t$.
3. Compute the sample variance

$$s^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Z_t - \bar{Z}_n)^2$$

4. We construct the 95% CI

$$\left[\overline{Z}_n - 1.96\sqrt{\frac{s^2}{n}}, \overline{Z}_n + 1.96\sqrt{\frac{s^2}{n}}\right]$$

5. We check whether 0 lies in that interval.
6. Repeat above steps for 1000 rounds; then report the empirical coverage, i.e., the proportion of times 0 is covered.

The empirical coverage probability (Method 1) = 0.738.

2.3 Using Sectioning Method

For the sectioning method, the algorthim would be as follows:

1. We generate $Z_1, \ldots, Z_n$ for $n = 500$ with $a = 0.5$.
2. We need to select batches to partition the series. Since we have n = 500 samples, we can derive in into L=10 segments of length B=50. ($L * B = n$)
3. Compute the **Batch Mean**

$$M_i = \frac{1}{B}\sum_{t=(i-1)\cdot B+1}^{i\cdot B} Z_t \quad \text{for } i = 1, \ldots, L$$

4. Calculate Overall Sample mean

$$\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i = \frac{1}{L}\sum_{i=1}^{L} M_i$$

5. Estimate the variance of $\bar{Z}_n$ using M_i

$$\widehat{Var}(\overline{Z}_n) = \frac{1}{L(L-1)} \sum_{i=1}^{L} (M_i - \overline{M})^2$$

6. The standard error would be

$$\hat{SE} = \sqrt{\frac{1}{L(L-1)} \sum_{i=1}^{L} (M_i - \overline{M})^2}$$

7. We construct the 95% CI
$$\left[\overline{Z}_n - 1.96\hat{SE}, \overline{Z}_n + 1.96\hat{SE}\right]$$

8. We check whether 0 lies in that interval.

9. Repeat above steps for 1000 rounds; then report the empirical coverage, i.e., the proportion of times 0 is covered.

Empirical coverage probability (Method 3, sectioning) = 0.918.

2.4 Interpretation

The return of the 3 methods are correspondingly 0.956, 0.738, 0.918.

Essentially, method 1, using the true theoretical variance would be the best case, since we have already know the exact auto-correlation. However, this would not happen in real life.

The return of method 2 is much smaller than the accuracy in method 1 and 2, mainly because the when we used the sample variance, we mistakes the autocorrelated data to be i.i.d. samples. The correlation inflates the variability of the sample mean since the estimation in method 2 ignores the correlation, the standard error would be smaller, thus narrower the intervals.

In method 3, we implement the batch-means method, which is an approach that is model-free. This approach highly reduces the time dependence by grouping observations together. This approach is more practical than method 2, because it does not require pre-knowledge for a. It would generates the CI more closer to 95%, but still not enough to reach the method using true variance.