Chapter 3 - Monte Carlo Methods

Monte Carlo applications to Statistical Inference.
Confidence Intervals.

Prof. Alex Alvarez, Ali Raisolsadat

School of Mathematical and Computational Sciences
University of Prince Edward Island

Definition: A confidence interval with confidence coefficient $1-\alpha$ for a parameter θ is a random interval $[U, V] \subset \mathbb{R}$ where $U = U(X)$ and $V = V(X)$ are functions of the random sample $X = (X_1, X_2, ..., X_n)$ such that:

$$P_\theta \left(\theta \in [U(X), V(X)]\right) \geq 1-\alpha$$

for all $\theta \in \Theta$. The subscript θ on the probability P indicates that the random sample X is assumed to be distributed according to the distribution with parameter θ

Remark: For the purpose of today's class, θ is a one-dimensional parameter. Some of these ideas can also be considered in multidimensional parametric spaces.

Confidence Intervals

In many cases we can construct (analytically) confidence intervals $[U(X), V(X)]$ that satisfy the definition above.

One well know example of that is the case in which $X_1, X_2, \ldots, X_n$ are i.i.d. with distribution $\sim N(\mu, \sigma^2)$ where the variance σ^2 is known.

Let $\hat{\mu}(X) = \frac{1}{n}\sum_{i=1}^{n} X_i$. The confidence interval for the parameter μ given by

$$U(X) = \hat{\mu}(X) - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \quad V(X) = \hat{\mu}(X) + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

satisfies that $P_\mu(\mu \in [U(X), V(X)]) = 1 - \alpha$.

We can do this because we know the exact distribution of our point estimator $\hat{\mu}(X)$.

Confidence Intervals

This interval estimator for the mean can be extended to other situations as long as n is large. We might need to replace σ with s (the standard deviation of the sample X). That is because for large values of n, the Central Limit Theorem applies.

In other situations, for instance if the data is not normally distributed, or if n is small it might be more complicated to construct confidence intervals.

Today we will see how we can use Monte Carlo simulations to assess whether a confidence interval $[U, V]$ is appropriate.

For simplicity we will still consider a confidence interval of the mean parameter of some distribution.

Example 3.38 Consider $X = (X_1, X_2, \ldots, X_n)$ i.i.d. Poisson distributed with unknown parameter $\lambda \in [0.1, 1]$. Assume that $n = 10$; meaning that Central Limit Theorem approximations do not apply.

Consider the approximated confidence interval for the mean parameter λ given by:

$$U(X) = \hat{\lambda}(X) - t^{(n-1)}_{1-\alpha/2} \frac{s}{\sqrt{n}} \quad V(X) = \hat{\lambda}(X) + t^{(n-1)}_{1-\alpha/2} \frac{s}{\sqrt{n}}$$

where $t^{(n-1)}_{1-\alpha/2}$ is the percentile at level $1 - \alpha/2$ of the Student's t distribution with n-1 degrees of freedom, and s is the sample standard deviation. Use $\alpha = 0.05$.

Problem: Do we actually have that

$$P_\lambda(U(X) \leq \lambda \leq V(X)) \geq 0.95$$

for all $\lambda \in [0.1, 1]$?

Algorithm:

1. For fixed λ generate samples $\{X^{(j)}\}_{j=1,2,\ldots N}$ of size 10 and compute $U^{(j)}$, $V^{(j)}$ for each sample.
2. Approximate
$$P_\lambda(U(X) \leq \lambda \leq V(X)) \approx \frac{1}{N}\sum_{j=1}^{N} 1_{\left[U^{(j)},V^{(j)}\right]}(\lambda)$$
3. Repeat steps 1 and 2 for as many values of λ as needed

Algorithm 1 Monte Carlo Estimation of Confidence Coefficients for Poisson Mean

Input: Number of samples n, number of simulations N, Poisson rate grid $\{\lambda_i\}_{i=1}^{K}$
for $i = 1$ to K **do**
  Initialize counter $k \leftarrow 0$
  for $j = 1$ to N **do**
    Generate $X_{j1}, \ldots, X_{jn} \sim \text{Poisson}(\lambda_i)$
    Compute sample mean $\bar{X}_j = \frac{1}{n}\sum_{t=1}^{n} X_{jt}$
    Compute standard deviation $s_j = \text{sd}(X_j)$
    Compute margin of error $d_j = s_j \cdot t_{0.975,\, n-1}/\sqrt{n}$
    if $\bar{X}_j - d_j \leq \lambda_i \leq \bar{X}_j + d_j$ **then**
      Increment counter: $k \leftarrow k + 1$
    end if
  end for
  Estimate coverage probability: $P_i = k/N$
end for
Output: Coverage probabilities $\{P_i\}_{i=1}^{K}$ as a function of λ_i

From Nested Loops to Vectorized Mathematics

Nested-loop formulation:

$$\text{For each } \lambda_i,\ i = 1, \ldots, K \quad \text{and each simulation } j = 1, \ldots, N:$$

$$X_{ij1}, \ldots, X_{ijn} \sim \text{Poisson}(\lambda_i)$$

$$\bar{X}_{ij} = \frac{1}{n}\sum_{t=1}^{n} X_{ijt}, \qquad s_{ij} = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(X_{ijt} - \bar{X}_{ij})^2}$$

$$d_{ij} = s_{ij} \cdot \frac{t_{0.975,\, n-1}}{\sqrt{n}}, \qquad I_{ij} = \begin{cases} 1, & \text{if } \bar{X}_{ij} - d_{ij} \le \lambda_i \le \bar{X}_{ij} + d_{ij} \\ 0, & \text{otherwise} \end{cases}$$

$$P_i = \frac{1}{N}\sum_{j=1}^{N} I_{ij}$$

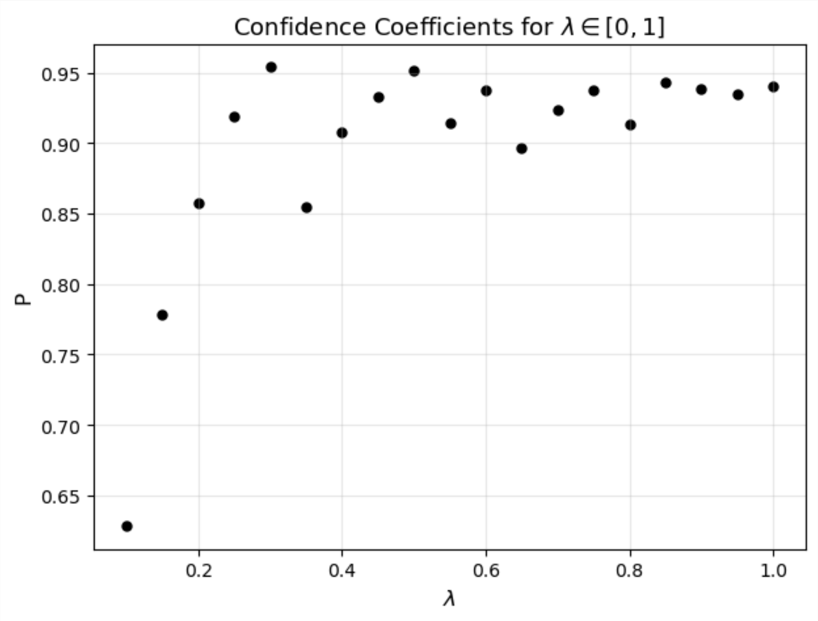
Matrix/vectorized representation:

$$\mathbf{X} = [X_{ijt}] \in \mathbb{R}^{K \times N \times n}, \quad X_{ijt} \sim \text{Poisson}(\lambda_i)$$

$$\bar{\mathbf{X}} = \text{mean}(\mathbf{X}, \text{axis} = 3), \quad \mathbf{S} = \text{sd}(\mathbf{X}, \text{axis} = 3)$$

$$\mathbf{D} = \mathbf{S} \cdot \frac{t_{0.975,\, n-1}}{\sqrt{n}}, \quad \boldsymbol{\Lambda} = [\lambda_i]_{i=1}^{K} \in \mathbb{R}^{K \times 1}$$

$$\mathbf{I} = (\bar{\mathbf{X}} - \mathbf{D} \le \boldsymbol{\Lambda} \le \bar{\mathbf{X}} + \mathbf{D}), \quad \mathbf{P} = \frac{1}{N}\sum_{j=1}^{N} \mathbf{I}_{ij}$$


Confidence Coefficients for $\lambda \in [0, 1]$
0.95
0.90
0.85
0.80
0.75
0.70
0.65
P
0.2
0.4
0.6
0.8
1.0
λ

Remarks

- We can see that for small values of λ, the confidence interval will contain the actual value of λ much less frequently than the desired 95% of the time. This indicates that proper 95% confidence intervals must be larger.
- A more detailed analysis would be needed in order to decide how to much increase should we increase that interval.
- Not necessarily the confidence interval has to be symmetric around the point estimator.

Homework

Perform an analysis similar to what we did in **Example 3.38** but now using samples of size $n = 8$ from the exponential distribution with mean $\mu \in [5, 10]$. Do you think that U and V as defined in **Example 3.38** can be used as an appropriate 95% confidence interval for $\mu \in [5, 10]$

Reminder: Random variables with exponential distributions with intensity λ have expected value $\mu = 1/\lambda$.