1. **Confidence intervals for normally distributed data** [4 Points]

10.1

Suppose that in an experiment 5 data points are observed:

$$x_1 = 2.3, \qquad x_2 = 1.9, \qquad x_3 = 2.0, \qquad x_4 = 1.8, \qquad x_5 = 2,1.$$

Suppose that these data points are realizations of i.i.d. random variables $X_1, \dots, X_5$, where $X_1 \sim \mathcal{N}(\mu, \sigma^2)$ with *unknown* μ and σ^2.

(a) Find a 95%-confidence interval for μ.

(a)

$$S(x) = \left[\overline{X}_n - \frac{S_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}, \ \overline{X}_n + \frac{S_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}} \right]$$

We input

$$\begin{cases} \overline{X}_n = 2.02 \quad t_{n-1,1-\frac{\alpha}{2}} = 2.78. \\ S_n^2 = 0.037 \Leftrightarrow S_n = \\ \alpha = 0.05 \\ n = 5 \end{cases} \Rightarrow S(x) = [1.781, \ 2.259]$$

(b) Explain why for general $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$ random variables, $\alpha \in (0,1)$, the interval

$$S(\mathbf{X}) = \left[\frac{n-1}{\chi^2_{n-1,1-\frac{\alpha}{2}}} S_n^2, \frac{n-1}{\chi^2_{n-1,\frac{\alpha}{2}}} S_n^2 \right], \qquad S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$$

is a $(1-\alpha)$-confidence interval for σ^2 when $\chi^2_{k,\beta}$ describes the β-quantile of the χ^2 distribution with k degrees of freedom. Calculate a 95%-confindence interval for σ^2 based on the observed data in the concrete example.

Hint: Use Theorem 2.7, (i).

proof. Thm 2.7 tells us

$$(n-1) \frac{S_n^2}{\sigma^2} \sim \chi^2_{n-1}$$

We are to calculate:
$$\mathbb{P}_{(\mu,\sigma^2)} \left[\sigma^2 \in \left[\frac{(n-1) S_n^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1) S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right] \right]$$

LHS
$$\sigma^2 \geq \frac{(n-1) S_n^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \Rightarrow \frac{(n-1) S_n^2}{\sigma^2} \leq \chi^2_{n-1,1-\frac{\alpha}{2}}$$

RHS.
$$\sigma^2 \leq \frac{(n-1) S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \Rightarrow \frac{(n-1) S_n^2}{\sigma^2} \geq \chi^2_{n-1,\frac{\alpha}{2}}$$

As $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$

$$\Rightarrow \mathbb{P}[\sigma^2 \in S(\mathbf{X})] = \mathbb{P}\left[\chi^2_{n-1,\frac{\alpha}{2}} \le \frac{(n-1)S_n^2}{\sigma^2} \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right]$$

$$= \mathbb{P}\left\{\chi^2_{n-1,\frac{\alpha}{2}} \le \chi^2_{n-1} \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right\}$$

$$= 1-\frac{\alpha}{2}-\frac{\alpha}{2} = 1-\alpha.$$

It follows that $S(x)$ is a confidence interval for σ^2.

- We use $S(x)$ to calculate a 95% confidence interval.

Input $\begin{cases} \alpha = 2.5\% = 0.025. \\ S_n^2 = 0.037 \\ n = 5 \\ \chi^2_{n-1,1-\frac{\alpha}{2}} = 11.143 \\ \chi^2_{n-1,\frac{\alpha}{2}} = 0.484 \\ \alpha = 0.025 \end{cases}$

$$\Rightarrow S(x) = [0.0133,\ 0.3058]$$

2. **Asymptotic confidence interval for the Poisson distribution** [4 points]

Let $X_1, ..., X_n \overset{\text{i.i.d.}}{\sim} Pois(\lambda)$, with unknown $\lambda > 0$.

(a) Calculate the Fisher information $I(\lambda)$.

(b) For $\alpha \in (0,1)$, find an asymptotic $(1-\alpha)$-confidence interval for λ.

(a) Fisher-Information

$$I(\lambda) = \mathbb{E}_\lambda \left[-\frac{d^2}{d\lambda^2} \log f_\theta(X_1) \Big|_{\theta=\lambda} \right]$$

$$f_\theta(X_1) = \frac{\lambda^{X_1} e^{-\lambda}}{X_1 !}$$

$$\log f_\theta(x_1) = X_1 \log\lambda - \lambda - \log X_1 !$$

$$-\frac{d^2}{d\lambda^2} \log f_\theta(x_1) = +X_1 \frac{1}{\lambda^2} = \frac{X_1}{\lambda^2}$$

$$\Rightarrow I(\lambda) = \mathbb{E}_\lambda \left[-\frac{d^2}{d\lambda^2} \log f_\theta(x_1) \Big|_{\theta=\lambda} \right] = \mathbb{E}_\lambda \left[\frac{X_1}{\lambda^2} \right] = \frac{1}{\lambda}$$

(b) First we need to compute a MLE estimator

$$\frac{d}{d\lambda} \sum \log f_\lambda(x_i) = \frac{d}{d\lambda} \sum X_i \log\lambda - \lambda - \log x_i !$$

$$\Rightarrow \frac{\sum x_i}{\lambda} - n = 0 \Rightarrow \hat{\lambda}_n = \frac{1}{n}\left(\sum_{i=1}^n x_i\right)$$

Then we verify the Convergence of Fisher Information.

$$I(\hat{\lambda}_n) = \frac{n}{\sum_{i=1}^n x_i}.$$ By Law of large numbers, $\sum_{i=1}^n x_i \xrightarrow{\mathbb{P}} \lambda$

with CMT, $I(\hat{\lambda}_n) \longrightarrow I(\lambda)$

Therefore, we may apply the Thm.

$$S(X) = \left[\hat{\theta}_n - \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n\, I(\hat{\theta}_n)}},\ \hat{\theta}_n + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n\, I(\hat{\theta}_n)}} \right]$$

where: $\begin{cases} \hat{\theta}_n = \dfrac{\sum_{i=1}^{n} X_n}{n} \\ I(\hat{\theta}_n) = \dfrac{n}{\sum_{i=1}^{n} X_n} \\ u_{1-\frac{\alpha}{2}} \text{ is the } 1-\frac{\alpha}{2} \text{ quantile of Normal distribution.} \end{cases}$

3. **Asymptotic confidence interval for the Exponential distribution** [4 points]

Let $X_1, ..., X_n \overset{\text{i.i.d.}}{\sim} \mathcal{E}(\theta)$, with unknown $\theta > 0$.

(a) Calculate the Fisher information $I(\theta)$.

(b) For $\alpha \in (0,1)$, find an asymptotic $(1-\alpha)$-confidence interval for θ.

(a)

$$\log f_\theta(X_1) = \log \theta e^{-\theta X_1} = \log\theta - \theta X_1$$

$$-\frac{d^2 \log f_\theta(x_1)}{d\theta^2} = \frac{1}{\theta^2}$$

$$\Rightarrow \mathbb{E}_\theta\left[-\frac{d^2 \log f_\theta(x_1)}{d\theta^2}\right] = \frac{1}{\theta^2}$$

(b) Similarly.

$$\text{MLE} := \frac{\partial}{\partial\theta}\sum_{i=1}^{n} f_\theta(x_i) = \frac{\partial}{\partial\theta}\sum_{i=1}^{n}\log\theta e^{-\theta x_i}$$

$$= \frac{\partial}{\partial\theta}\sum \log\theta - \theta x_i$$

$$= \sum_{i=1}^{n}\frac{1}{\theta} - x_i = 0 \quad \Rightarrow \hat{\theta}_n = \frac{n}{\sum x_i}$$

$I(\hat{\theta}_n) \xrightarrow{n} I(\theta)$ can be verified

Therefore. the confidence interval is.

$$S(X) = \left[\hat{\theta}_n - \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n\, I(\hat{\theta}_n)}},\ \hat{\theta}_n + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n\, I(\hat{\theta}_n)}}\right]$$

where
$$\begin{cases} \hat{\theta}_n = \dfrac{n}{\sum_{i=1}^{n} x_i} \\ I(\hat{\theta}_n) = \dfrac{1}{\hat{\theta}_n^2} \end{cases}$$

$u_{1-\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$ - quantile for Normal distribution.

4. **Exact confidence interval for the Exponential distribution** [4 points]

Let $X_1, ..., X_n \overset{\text{i.i.d.}}{\sim} \mathcal{E}(\theta)$, with unknown $\theta > 0$.

(a) Show that $2\theta \sum_{i=1}^{n} X_i$ follows a χ^2_{2n} distribution.

(b) Use (a) to find an *exact* $(1-\alpha)$-confidence interval for θ, for $\alpha \in (0,1)$.

?

(a) Proof: $X_i \sim \mathcal{E}(\theta) \sim \Gamma(1, \theta)$

$$\Rightarrow \sum_{i=1}^{n} X_i \sim \Gamma(n, \theta)$$

$$\Rightarrow 2\theta \sum_{i=1}^{n} X_i \sim \Gamma\left(n, \frac{\theta}{2\theta}\right) = \Gamma\left(n, \frac{1}{2}\right)$$

The density of $\Gamma(n, \frac{1}{2})$ is

$$f(x) = \frac{(\frac{1}{2})^n}{\Gamma(n)} x^{n-1} e^{-\frac{1}{2}x} \mathbb{1}_{[0,+\infty)}(x)$$

The density of χ^2_{2n} is

$$f_{\chi^2_{2n}}(x) = \frac{1}{2^n \Gamma(n)} x^{n-1} e^{-\frac{x}{2}} \mathbb{1}_{[0,+\infty)}(x).$$

which coincides with the law of $\Gamma(n, \frac{1}{2})$

Therefore, $2\theta \sum_{i=1}^{n} X_i \sim \Gamma(n, \frac{1}{2}) \sim \chi^2_{2n}$

(b) $2\theta \sum_{i=1}^{n} x_i \sim \chi^2_{2n}$ as a result.

Let $u_{\frac{\alpha}{2}}$, $u_{1-\frac{\alpha}{2}}$ be $\frac{\alpha}{2}$, $1-\frac{\alpha}{2}$ quantile of χ^2_{2n}

we claim

$$\left[\frac{u_{\frac{\alpha}{2}}}{2\sum_{i=1}^{n} x_i}, \frac{u_{1-\frac{\alpha}{2}}}{2\sum_{i=1}^{n} x_i} \right]$$ is a $1-\alpha$ confidence interval of θ.

Indeed.

$$\mathbb{P}\left[\theta \in \left[\frac{u_{\frac{\alpha}{2}}}{2\sum_{i=1}^{n} x_i}, \frac{u_{1-\frac{\alpha}{2}}}{2\sum_{i=1}^{n} x_i}\right]\right]$$

$$= \mathbb{P}\left[u_{\frac{\alpha}{2}} \le 2\sum_{i=1}^{n} x_i \theta \le u_{1-\frac{\alpha}{2}} \right]$$

$$= (1-\tfrac{\alpha}{2}) - \tfrac{\alpha}{2} = 1-\alpha$$