(1) **Exponential family**
Show that the one-parameter exponential family has a monotone likelihood ratio in a sufficient statistic $T(\mathbf{X})$ if the natural parameter $w(\theta)$ is a non-decreasing function in θ.

The pdf is $f_\theta(x) = h(x)\, c(\theta)\, e^{w(\theta)t(x)}$, where $h, c \geq 0$, hence, the likelihood is

$L(\theta; x) = (c(\theta))^n \exp\left(w(\theta) \sum_{i=1}^n t(x_i)\right) \prod_{i=1}^n h(x_i)$. We obtain the likelihood ratio

$$\lambda(x) = \frac{L(\theta_1; x)}{L(\theta_2; x)} = \left(\frac{c(\theta_1)}{c(\theta_0)}\right)^n \exp\left((w(\theta_1) - w(\theta_0)) \sum_{i=1}^n t(x_i)\right).$$

The statistic $T(X) = \sum_{i=1}^n t(X_i)$ is sufficient by lecture 8 slide 48. Since w is non-decreasing, we have $w(\theta_1) - w(\theta_0) \geq 0$, because $\theta_1 > \theta_0$. Therefore, λ is a non-decreasing function of $T(x)$.

(2) **Confidence interval 1**
In the June 1986 issue of Consumer Reports, some data on the calorie content of beef hot dogs is given. Here are the numbers of calories in 20 different hot dog brands:

$$186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131, 149, 135, 132.$$

Assume that the numbers are from a normal distribution with mean μ and variance σ^2, both unknown. Use `R` to obtain a 90% confidence interval for the mean number of calories μ.

We can estimate the mean $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$ and the variance $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

The confidence interval is given by $[\hat{\mu} - \delta, \hat{\mu} + \delta)$, where $\mathbb{P}(\hat{\mu} - \delta \leq X < \hat{\mu} + \delta) = \alpha$ with $X \sim \mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$ and $1 - \alpha = \frac{9}{10}$ or $\alpha = \frac{1}{10}$. We have

$$1 - \alpha = \mathbb{P}(\hat{\mu} - \delta \leq X < \hat{\mu} + \delta) = 1 - 2\mathbb{P}(X < \hat{\mu} - \delta) = 1 - 2\mathbb{P}\left(\frac{(X - \hat{\mu})\sqrt{n}}{\hat{\sigma}} < -\frac{\delta\sqrt{n}}{\hat{\sigma}}\right)$$

$$\Leftrightarrow \phi\left(-\frac{\delta\sqrt{n}}{\hat{\sigma}}\right) = \frac{\alpha}{2} \Leftrightarrow -\frac{\delta\sqrt{n}}{\hat{\sigma}} = \phi^{-1}\left(\frac{\alpha}{2}\right) \Leftrightarrow \delta = -\frac{\hat{\sigma}}{\sqrt{n}}\phi^{-1}\left(\frac{\alpha}{2}\right)$$

Using R, we obtain $\hat{\mu} \approx 156{,}85$ and $\hat{\sigma} \approx 22{,}642$ and $\delta \approx 8{,}328$

Hence, $\hat{\mu} - \delta \approx 148{,}522$ and $\hat{\mu} + \delta \approx 165{,}178$

(3) **Confidence interval 2**
Suppose $X_1, \dots, X_n$ are i.i.d. with pdf

$$f(x|\lambda, \eta) = \begin{cases} \lambda e^{-\lambda(x-\eta)} & x > \eta \\ 0, & \text{otherwise} \end{cases}$$

where λ and η positive parameters with η known but λ unknown. Find the MLE of λ and construct a $(1-\alpha)100\%$ confidence interval for λ when n is assumed to be large.

The likelihood function is

$$L(\lambda, \eta; x) = \begin{cases} \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} (x_i - \eta)\right), & \text{if } \min\{x_i \mid 1 \le i \le n\} > \eta \\ 0 & , \text{otherwise} \end{cases}$$

Thus, the log-likelihood for $x \in (\eta, \infty)^n$ is $\ell(\lambda, \eta; x) = n\log(\lambda) - \lambda \sum_{i=1}^{n} (x_i - \eta)$ and

$$\frac{\partial \ell}{\partial \lambda}(\hat{\lambda}, \eta; x) = \frac{n}{\hat{\lambda}} - \sum_{i=1}^{n} (x_i - \eta) \overset{!}{=} 0 \Leftrightarrow \hat{\lambda} = n\left(\sum_{i=1}^{n} (x_i - \eta)\right)^{-1} \text{ is the MLE of } \lambda.$$

$$\frac{\partial^2 \ell}{\partial \lambda}(\lambda, \eta; x) = -\frac{n}{\lambda^2} < 0 \quad \text{Hence}, \quad I_n(\lambda) = -\mathbb{E}\left(\frac{\partial^2 \ell}{\partial \lambda}(\lambda, \eta; x)\right) = \frac{n}{\lambda^2}$$

By slide 12 from lecture 7 we have $\sqrt{n}\,(\hat{\lambda}(X) - \lambda) \xrightarrow{d} \mathcal{N}\left(0, \frac{\lambda^2}{n}\right)$, hence the MLE is approximately $\mathcal{N}\left(\hat{\lambda}, \frac{\hat{\lambda}^2}{n}\right)$ distributed.

Let $Z \sim \mathcal{N}\left(\hat{\lambda}, \frac{\hat{\lambda}^2}{n}\right)$. The confidence interval $[\hat{\lambda} - \delta, \hat{\lambda} + \delta)$ is such that

$$1 - \alpha = \mathbb{P}(\hat{\lambda} - \delta \le Z < \hat{\lambda} + \delta) = 1 - 2\,\mathbb{P}(Z < \hat{\lambda} - \delta) = 1 - 2\,\mathbb{P}\left(\frac{(Z - \hat{\lambda})\sqrt{n}}{\hat{\lambda}} < -\frac{\delta\sqrt{n}}{\hat{\lambda}}\right)$$

$$\Leftrightarrow \phi\left(-\frac{\delta\sqrt{n}}{\hat{\lambda}}\right) = \frac{\alpha}{2} \Leftrightarrow -\frac{\delta\sqrt{n}}{\hat{\lambda}} = \phi^{-1}\left(\frac{\alpha}{2}\right) \Leftrightarrow \delta = -\frac{\hat{\lambda}}{\sqrt{n}}\,\phi^{-1}\left(\frac{\alpha}{2}\right)$$

(4) **Confidence inerval 3**
Use R to generate a random sample $X_1, \dots, X_n$ from $Pois(1)$ distribution (for $n = 30$ and $n = 100$). Compute the 90% condidence interval for λ, check if it contains the true value of $\lambda = 1$, and repeat this 10000 times. What is the fraction of simulations for which the confidence interval covers λ?

We saw in lecture 7 on slide 16 that the MLE is $\hat{\lambda} = \bar{X}$ with Variance $\frac{\hat{\lambda}}{n}$. So we consider $1-\alpha = \mathbb{P}(\hat{\lambda} - \delta \leq Z < \hat{\lambda} + \delta) = \mathbb{P}(-\delta \leq Z - \hat{\lambda} < \delta) = \mathbb{P}\left(-\frac{\delta\sqrt{n}}{\sqrt{\hat{\lambda}}} \leq \frac{\sqrt{n}(Z-\hat{\lambda})}{\sqrt{\hat{\lambda}}} < \frac{\delta\sqrt{n}}{\sqrt{\hat{\lambda}}}\right)$

$$= 1 - 2\phi\left(-\frac{\delta\sqrt{n}}{\sqrt{\hat{\lambda}}}\right) \Leftrightarrow \phi\left(-\frac{\delta\sqrt{n}}{\sqrt{\hat{\lambda}}}\right) = \frac{\alpha}{2} \Leftrightarrow \delta = -\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}}\phi^{-1}\left(\frac{\alpha}{2}\right).$$

$\alpha = \frac{1}{10}$

(5) **Boxplot and quantiles**
Two novel randomized algorithms (A and B) are to be compared regarding their runtimes. Both algorithms were executed n times. The runtimes (in seconds) are stored in the file `algorithms.Rdata`

(a) Set the working directory and load the data using `load()`. Create a boxplot to compare the running times. Color the boxes and add proper notations (axes notations, title etc.). More info via `?boxplot`

(b) Comment on the following statements / questions only using the graphic
   (a) The first quartile of the times in A was about?
   (b) the interquartile range of the times in B is about trice the interquartile range of A
   (c) Is $n = 100$?
   (d) More than half of the running times in B were faster than 3/4 of the running times in A.
   (e) At least 50% in A were faster than the 25% slowest in B.
   (f) At least 60% in A were faster than the 25% slowest in B .

a) 20

b) for A: ≈ 12 for B: ≈ 24

c) can't tell

d) yes

e) yes

f) can't tell

(c) Regarding the runtimes

$$23.7,\ 13.7,\ 7.6,\ 9.0,\ 44.3,\ 3.5,\ 2.2,\ 34.2$$

which are a subset of B, find all empirical (a) medians, (b) first quartiles and (c) 2/3-quantiles (not using R).

2.2, 3.5, 7.6, 9.0, 13.7, 23.7, 34.2, 44.3

first quantiles [3.5, 7,6]

medians [9.0, 13.7]

2/3 - quantile: 23.7

$\mathbb{P}(X < x) \leq \frac{1}{4}$

$\mathbb{P}(X \leq x) \geq \frac{1}{4}$