Exercise H2.1. Let $X_1, \ldots, X_n$ be independent and identically distributed with Poisson law $\mathrm{Po}(\lambda)$, where $\lambda > \varepsilon$ is unknown, (ε is known, $0 < \varepsilon < n^{-1}$). . Find the maximum likelihood estimator (MLE) of λ (proof).

Exercise H2.2 Let $X_1, \ldots, X_n$ be independent and identically distributed such that X_1 has the uniform law on the set $\{1, \ldots, r\}$ for some integer $r > 1$ (i.e. $P(X_1 = k) = 1/r$, $k = 1, \ldots, r$). In the statistical model where $r > 1$ is unknown, find the MLE of r (proof).

Exercise H2.3. Let $X_1, \ldots, X_n$ be independent and identically distributed such that X_1 has the geometric law $\mathrm{Geom}(p)$, i.e.

$$P(X_1 = k) = (1-p)^{k-1}\, p,\ k = 1, 2, \ldots$$

In the statistical model where $p \in (0, \delta)$ is unknown (δ is known, $(1 + n^{-1})^{-1} < \delta < 1$) find the MLE of p (proof).

Exercise H2.4 This exercise presupposes the **continuous likelihood principle**. Let X be a random variable with values in $\mathbb{R}^k$ such that $\mathcal{L}(X) \in \{P_\vartheta; \vartheta \in \Theta\}$, and each law P_ϑ has a density $p_\vartheta(x)$ on $\mathbb{R}^k$. For each $x \in \mathbb{R}^k$, the function

$$L_x(\vartheta) = p_\vartheta(x)$$

is called the likelihood function of ϑ given x. A maximum likelihood estimator of ϑ is an estimator $T(x) = T_{ML}(x)$ such that

$$L_x(T_{ML}(x)) = \max_{\vartheta \in \Theta} L_x(\vartheta),$$

i.e. for each given x, the estimator is a value of ϑ which maximizes the likelihood.
Let $X_1, \ldots, X_n$ be independent and identically distributed such that X_1 has the normal law $N(\mu, \sigma^2)$. In the statistical model where $\mu \in \mathbb{R}$ is unknown and $\sigma^2 > 0$ is known, find the MLE of μ. (proof).

Exercise H2.5 This exercise refers to section 2.6. handout (conditional and posterior densities).Consider reversing the roles of ϑ and X, i.e. take the marginal probability function for X given by (2.29) and combine it with the conditional density for ϑ given by (2.30). Consider the expression $q_x(\vartheta)P(X = x)$ and, analogously to (2.31), divide it by its sum over all possible values of x ($x \in \mathcal{X}$). Call the result $q_\vartheta(x)$. Show that for any ϑ with $g(\vartheta) > 0$, the relation

$$q_\vartheta(x) = P_\vartheta(x),\ x \in \mathcal{X}$$

holds.
Remark. *This result justifies to call $P_\vartheta(x)$ a conditional probability function under $U = \vartheta$:*

$$P_\vartheta(x) = P\left(X = x | U = \vartheta\right),$$

even though U is continuous and the event $U = \vartheta$ has probability 0.

Exercise H2.1. Let $X_1,...,X_n$ be idependent annd identically distributed with Poisson law $Po\left(\lambda\right)$, where $\lambda>\varepsilon$ is unknown, (ε is known, $0<\varepsilon<n^{-1}$). . Find the maximum likelihood estimator (MLE) of λ (proof).

Solution: since $X_1,...,X_n$ are independent and identically distributed with Poisson law $Po\left(\lambda\right)$, then

$$L_x\left(\lambda\right)=P_\lambda\left(X_1=x_1,...,X_n=x_n\right)$$

$$=\prod_{i=1}^n P_\lambda\left(X_i=x_i\right)$$

$$=\prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!}e^{-\lambda}$$

$$=\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}e^{-n\lambda}.$$

Note that $\frac{d}{d\lambda}\left(\ln L_x\left(\lambda\right)\right)=\frac{\sum_{i=1}^n x_i}{\lambda}-n=0$ iff $\lambda=\frac{\sum_{i=1}^n x_i}{n}$, and $\frac{d^2}{d\lambda^2}\left(\ln L_x\left(\lambda\right)\right)=-\frac{\sum_{i=1}^n x_i}{\lambda^2}\leqslant 0$.

If $\sum_{i=1}^n x_i\geqslant 1$, the function $\ln L_x\left(\lambda\right)$ is strictly concave and achieves its maximum value at $\lambda=\frac{\sum_{i=1}^n x_i}{n}\left(>\varepsilon\right)$. Thus the MLE of λ is $\frac{\sum_{i=1}^n x_i}{n}$.

If $\sum_{i=1}^n x_i=0$, $\frac{d}{d\lambda}\left(\ln L_x\left(\lambda\right)\right)=-n<0$, then $\sup_\lambda\left(\ln L_x\left(\lambda\right)\right)=\lim_{\lambda\to\varepsilon}\ln L_x\left(\lambda\right)$, but it is unattainable on the set$(\varepsilon,+\infty)$. Thus the MLE of λ doesn't exist.

Exercise H2.2 Let $X_1,\ldots,X_n$ be independent and identically distributed such that X_1 has the uniform law on the set $\{1,\ldots,r\}$ for some integer $r>1$ (i.e. $P(X_1=k)=1/r$, $k=1,\ldots,r$). In the statistical model where $r>1$ is unknown, find the MLE of r (proof).

Solution: Since $X_1,...,X_n$ are independent and identically distributed according to the uniform distribution $U\left\{1,...,r\right\}$ for some integer $r\geqslant 2$, then

$$L_x\left(r\right)=P_r\left(X_1=x_1,...,X_n=x_n\right)$$

$$=\prod_{i=1}^n P_r\left(X_i=x_i\right)$$

$$=\frac{1}{r^n}1_{\{\max(x_1,...,x_n)\leqslant r\}}\left(x_1,...,x_n\right).$$

Note that $L_x\left(r\right)=\frac{1}{r^n}$ is decreaing on the set $\left\{r|r\geqslant\max\left\{x_1,...,x_n,2\right\}\right\}$.

This implies $L_x(r)$ achieves its maximum at $r = \max\{x_1, ..., x_n, 2\}$. Thus the MLE of r is $\max\{x_1, ..., x_n, 2\}$.

Exercise H2.3. Let $X_1, \ldots, X_n$ be independent and identically distributed such that X_1 has the geometric law Geom(p), i.e.

$$P(X_1 = k) = (1-p)^{k-1}\, p,\ k = 1, 2, \ldots$$

In the statistical model where $p \in (0, \delta)$ is unknown (δ is known, $(1 + n^{-1})^{-1} < \delta < 1$) find the MLE of p (proof).

Solution: since $X_1, .., X_n$ are independent and identically distributed according to the geometric law $Geom\,(p)$, then

$$\begin{aligned} L_x(p) &= P_r(X_1 = x_1, ..., X_n = x_n) \\ &= \textstyle\prod_{i=1}^n P_r(X_i = x_i) \\ &= \textstyle\prod_{i=1}^n (1-p)^{x_i - 1}\, p \\ &= (1-p)^{(\sum_{i=1}^n x_i) - n}\, p^n. \end{aligned}$$

Note that $\dfrac{d}{dp}(\ln L_x(p)) = -\dfrac{(\sum_{i=1}^n x_i) - n}{1-p} + \dfrac{n}{p} = \dfrac{n - p\sum_{i=1}^n x_i}{(1-p)\,p} = 0$ iff $p = \dfrac{n}{\sum_{i=1}^n x_i}$, and $\dfrac{d^2}{dp^2}(\ln L_x(p)) = -\dfrac{(\sum_{i=1}^n x_i) - n}{(1-p)^2} - \dfrac{n}{p^2} < 0.$

If $\sum_{i=1}^n x_i \geqslant n + 1$, then $\ln L_x(p)$ is strictly concave and achieves its maximum value at $p = \dfrac{n}{\sum_{i=1}^n x_i}\,(< \delta)$. Thus the MLE of p is $\dfrac{n}{\sum_{i=1}^n x_i}$.

If $\sum_{i=1}^n x_i = n$, then $\dfrac{d}{dp}(L_x(p)) > 0$, then $\sup_p L_x(p) = \lim_{p \to \delta} L_x(p)$, but it is unattainable on the set $(0, \delta)$. Thus the MLE of p doesn't exist.

Exercise H2.4 This exercise presupposes the **continuous likelihood principle**. Let X be a random variable with values in $\mathbb{R}^k$ such that $\mathcal{L}(X) \in \{P_\vartheta; \vartheta \in \Theta\}$, and each law P_ϑ has a density $p_\vartheta(x)$ on $\mathbb{R}^k$. For each $x \in \mathbb{R}^k$, the function

$$L_x(\vartheta) = p_\vartheta(x)$$

is called the likelihood function of ϑ given x. A maximum likelihood estimator of ϑ is an estimator $T(x) = T_{ML}(x)$ such that

$$L_x(T_{ML}(x)) = \max_{\vartheta \in \Theta} L_x(\vartheta),$$

i.e. for each given x, the estimator is a value of ϑ which maximizes the likelihood.
Let $X_1, \ldots, X_n$ be independent and identically distributed such that X_1 has the normal law $N(\mu, \sigma^2)$. In the statistical model where $\mu \in \mathbb{R}$ is unknown and $\sigma^2 > 0$ is known, find the MLE of μ. (proof).

Solution: since $X_1, ..., X_n$ are independent and identically distributed according the normal law $N(\mu, \sigma^2)$, then

$$\begin{aligned} L_x(\mu) &= P_\theta(X_1 = x_1, ..., X_n = x_n) \\ &= \prod_{i=1}^n P_\theta(X_i = x_i) \\ &= \prod_{i=1}^n \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}. \end{aligned}$$

Note that $\frac{d}{d\mu}\ln L_x(\mu) = -\frac{1}{2\sigma^2}\sum_{i=1}^n(-2(x_i-\mu)) = \frac{1}{\sigma^2}\left(\sum_{i=1}^n x_i - n\mu\right) = 0$ iff $\mu = \frac{\sum_{i=1}^n x_i}{n}$, and $\frac{d^2}{d\mu^2}\ln L_x(\mu) = -\frac{n}{\sigma^2} < 0$.

This implies $\ln L_x(\mu)$ achieves its maximum at $\mu = \frac{\sum_{i=1}^n x_i}{n}$. Thus the MLE of μ is $\frac{\sum_{i=1}^n x_i}{n}$.

Exercise H2.5 This exercise refers to section 2.6. handout (conditional and posterior densities).Consider reversing the roles of ϑ and X, i.e. take the marginal probability function for X given by (2.29) and combine it with the conditional density for ϑ given by (2.30). Consider the expression $q_x(\vartheta)P(X = x)$ and, analogously to (2.31), divide it by its sum over all possible values of x ($x \in \mathcal{X}$). Call the result $q_\vartheta(x)$. Show that for any ϑ with $g(\vartheta) > 0$, the relation

$$q_\vartheta(x) = P_\vartheta(x),\ x \in \mathcal{X}$$

holds.
Remark. *This result justifies to call $P_\vartheta(x)$ a conditional probability function under $U = \vartheta$:*

$$P_\vartheta(x) = P(X = x|U = \vartheta),$$

even though U is continuous and the event $U = \vartheta$ has probability 0.

Solution: for any θ with $g(\theta) \neq 0$ and x, $\acute{x}$, we have

$$q_\theta(x) = \frac{q_x(\theta)P(X = x)}{\sum_{x'\in\chi} q_{x'}(\theta)P(X = x')}.$$

From the definition of $q_x(\theta)$, we have

$$q_x(\theta) = \frac{P_\theta(x)g(\theta)}{P(X = x)},$$

then

$$q_x(\theta)P(X = x) = P_\theta(x)g(\theta),$$

and

$$\begin{aligned} \sum_{x'\in\chi} q_{x'}(\theta)P(X = x') &= \sum_{x'\in\chi} P_\theta(x')g(\theta) \\ &= g(\theta)\sum_{x'\in\chi} P_\theta(x') \end{aligned}$$

$$= g\left(\theta\right).$$

Thus $q_\theta\left(x\right) = P_\theta\left(x\right).$