Ch.2-Ex.7: (a) Since $X_i \sim \text{Poi}(\lambda)$ for $i = 1, \ldots, n$, and $\Pr(X_i = x_i) = \frac{\lambda^{x_i}}{x_i!}e^{-\lambda}$, the likelihood function is given by

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

(b) Notice that

$$\frac{l(\lambda_1)}{l(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^{\sum_{i=1}^{n} x_i} \exp\{-n(\lambda_1 - \lambda_2)\}$$

Hence, the likelihood function depends on $\lambda$ only through $\sum_{i=1}^{n} x_i$ (and $n$).

(c) Since the value of $\lambda$ that maximizes $l(\lambda)$ also maximizes $\log(l(\lambda))$ it is convenient to work with log-likelihood function:

$$\ln(l(\lambda)) = \sum_{i=1}^{n} x_i \lambda - n\lambda - \ln(\prod_{i=1}^{n} x_i!$$

$$\frac{d\ln(l(\lambda))}{d\lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$$

Equating to 0 yields

$$0 = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$$

$$\lambda = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Consider $\frac{d^2\ln(l(\lambda))}{d\lambda^2} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}$. When $\lambda = \bar{x}$, $\frac{d^2\ln(l(\lambda))}{d\lambda^2} = -\frac{n}{\bar{x}}$. If $\bar{x} \geq 0$, $\frac{d^2\ln(l(\lambda))}{d\lambda^2} \leq 0$. So at this time $\hat{\lambda} = \bar{y}$ is the m.l.e.

(d) In example 1.4, there is only one observation, $X = 3$. So the MLE of $\lambda$ is 3.
Ch.2-Ex.9: In example 2.8

\[ l(\theta) \propto \theta^8(1 - \theta)^{137} \]

Since \(\ln(x)\) is an increasing function in terms of \(x\), \(\theta\) which maximizes \(l(\theta)\) also maximizes \(\ln(l(\theta))\). Hence, to find out m.l.e., we can just consider the value of \(\theta\) that maximizes \(\ln(l(\theta))\).

\[
\ln(l(\theta)) \propto 8 \ln \theta 137 \ln(1 - \theta)
\]

\[
\frac{d\ln(l(\theta))}{d\theta} = \frac{8}{\theta} - \frac{137}{1 - \theta}
\]

Equating to 0 yeilds

\[
0 = \frac{8}{\theta} - \frac{137}{1 - \theta}
\]

\[
\theta = \frac{8}{145}
\]

Consider \(\frac{d^2\ln(l(\theta))}{d\theta^2} = -\frac{8}{\theta^2} + \frac{137}{(1 - \theta)^2}\). When \(\theta = \frac{8}{145}\), \(\frac{d^2\ln(l(\theta))}{d\theta^2} = -17.06 < 0\) So the m.l.e. for \(\theta\) is \(\frac{8}{145}\).

Ch.2-Ex.10*Method 1

Suppose that the exact distance between John’s office and home is \(\mu\). Then it is reasonable to assume that the distance \((X)\) John’s runs follows a distribution with mean \(\mu\). For simplicity we assume that \(X \sim N(\mu, 1)\), i.e., a normal distribution with (unknown) mean \(\mu\) and variance 1. A more flexible model would probably assume that the variance is also unknown.

Under the above assumption, the number displayed on the odometer \(Y\) has the probability \(\Pr(Y = d) = FX(d+1) - FX(d)\), where \(FX(x)\) denotes the cdf of \(X\). Notice that \(Y = [X] = \) the largest integer less than or equal to \(X\) (the “box” function). In the exercise, 3 is observed 7 times, and 4 is observed 3 times. Since the 10 observations are mutually independent, the likelihood function

\[
l(\mu) = \Pr(Y = 3)^7 \Pr(Y = 4)^3 = \left(\int_3^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx\right)^7 \left(\int_4^5 \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx\right)^3
\]

\[
\ln(l(\mu)) = 7 \ln \left(\int_3^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx\right) + 3 \ln \left(\int_4^5 \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx\right)
\]

To find the m.l.e., we can use R programming:

```r
> loglik=function(mu,sigma=1){3*log(pnorm(4,mu,sigma)-pnorm(3,mu,sigma)) + 7*log(pnorm(5,mu,sigma)-pnorm(4,mu,sigma))}
> curve(loglik,from=2,to=6)
> optimize(loglik,c(3,5),maximum=T,sigma=1)
```

©Sujit Ghosh, NCSU Statistics
$maximum$
\[1\] 4.1998

$objective$
\[1\] -10.56479

>abline(v=4.1998)

So the m.l.e. of the exact distance is 4.1998.

method 2 Suppose the number displayed on the odometer $Y \sim Possion(\lambda)$. $\lambda$ is the exact distance between John’s office and home.
Using the result of Ch2.Ex7, we can get the estimate of $\lambda = \bar{y} = \frac{3\times7 + 4\times3}{10} = 3.3$.

method 3 The measure of distance between home and office should be a continuous variable. One generally used continuous distribution is normal distribution.
We can assume that $Y \sim N(\mu, \sigma)$. Here, it is easy to see that $\mu$ represents the exact distance between home and office.

\[
\log(l(\mu, \sigma)) = \prod_{i=1}^{10} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_i - \mu)^2}{2\sigma^2}}
\]

Use the result of Ch2.Ex.13, we can get the estimate of $\hat{\mu} = \bar{y} = 3.3$

Using methods 2 and 3, we will always get an estimate between 3 and 4. But the exact distance exceeds 4 when the odometer shows 4. This is a drawback when you use methods 2 and 3. For method 1, I assume the variance to be 1, which may not be reasonable.
However, this assumption dose not affect the estimate of \( \mu \) much. Even setting the variance to be 10, the estimate almost stayed the same. For example, try the following R commands:

```r
for(sigma in seq(0.5,2.5,l=10)){
cat(c("sigma=",round(sigma,3)),fill=T)
print(optimize(loglik,c(2.5,5.5),maximum=T,sigma=sigma))
}
```

Ch.2-Ex.11 The m.l.e \( \hat{\mu} \) is obtained from the equation (b).

Ch.2-Ex.12 (a) Since \( Y_i \overset{i.i.d}{\sim} N(\mu,1) \) and \( f_{Y_i}(y_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_i-\mu)^2}{2}} \), the likelihood function is given by

\[
l(\mu) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \left( \frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2}}
\]

(b) Notice that

\[
\frac{l(\mu_1)}{l(\mu_2)} = \exp\left\{-\frac{1}{2} \left[ \sum_{i} (y_i - \mu_1)^2 - \sum_{i} (y_i - \mu_2)^2 \right] \right\}
\]

\[
= \exp\left\{-\frac{1}{2} \left[ 2(\mu_2 - \mu_1) \sum_{i} y_i + n(\mu_1^2 - \mu_2^2) \right] \right\}
\]

We can see that in the above expression depends only on \( \sum_{i=1}^{n} y_i \) (and \( n \)).

(c) Set \( \mu = 0 \) for illustration.

R codes:

```r
mu0=0;n=10
y=rnorm(n,mean=mu0)
lik=function(mu){prod(dnorm(y,mu))/prod(dnorm(y,mean(y)))}
likelihood=function(mu){sapply(mu,lik)}
mu=seq(-3,3,l=100)
plot(mu,likelihood(mu),type="l")
for(l in 1:30){
y=rnorm(n,mean=mu0)
lines(mu,likelihood(mu),col="gray")
}
```
We need to find the likelihood set \( LS_{0.1} = \{ \mu : l(\mu) \geq 0.1l(\hat{\mu}) \} \) to characterize the accuracy of estimation. It can be shown that for this example,

\[
LS_\alpha = \left[ \bar{y} + \frac{2}{n} \log(\alpha), \bar{y} - \frac{2}{n} \log(\alpha) \right] \quad \text{for any } \alpha \in (0,1).
\]

In particular, for \( \alpha = 0.1 \) and \( n = 10 \), the length of the interval is 0.921.

(d) Since \( Y_i \overset{i.i.d.}{\sim} N(\mu, \sigma) \) and \( f_{Y_i}(y_i) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}} \) the likelihood is given by

\[
l(\mu) = \prod_{i=1}^{10} f_{Y_i}(y_i) = \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^{10} e^{-\frac{\sum_{i=1}^{10}(y_i-\mu)^2}{2\sigma^2}}
\]

Here, \( \sigma \) is assumed known.

(e) Similar to (d)

\[
l(\sigma) = \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^{10} e^{-\frac{\sum_{i=1}^{10}(y_i-\mu)^2}{2\sigma^2}}
\]

Here, \( \mu \) is assumed known.

Ch.2-Ex.13 Use the result of Ch.2-Ex.12(a),

\[
l(\mu) \propto e^{-\frac{-2\sigma \sum_{i=1}^{n} y_i + n\mu^2}{2}}
\]

Since \( \ln(x) \) is an increasing function in terms of \( x \), \( \mu \) which maximizes \( l(\mu) \) also maximizes \( \ln(l(\mu)) \). Hence, to find out m.l.e., we can just consider \( \ln(l(\mu)) \). It easily follows that \( \frac{d\ln(l(\mu))}{d\mu} = \sum_i y_i - n\mu = 0 \) if \( \mu = \bar{y} = \sum_i y_i/n \). Also it follows easily that \( \frac{d^2\ln(l(\mu))}{d\mu^2} = -n < 0 \) and hence the MLE of \( \mu \) is \( \bar{y} \).
Ch.2-Ex.14 Use the result of Ch.2-Ex.12(e),

\[ l(\sigma) = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}} \propto \sigma^{-n} e^{-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}} \]

Since \( \ln(x) \) is an increasing function in terms of \( x \), \( \sigma \) which maximizes \( l(\sigma) \) also maximizes \( \ln(l(\sigma)) \). Hence, to find out m.l.e., we can just consider \( \ln(l(\sigma)) \). It follows easily that

\[
\frac{d\log(l(\sigma))}{d\sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{\sigma^3}.
\]

Equating to 0 yeilds

\[
0 = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{\sigma^3} \Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{n}.
\]

Consider

\[
\frac{d^2\ln(l(\sigma))}{d\sigma^2} = \frac{n}{\sigma^2} - \frac{3\sum_{i=1}^{n}(y_i-\bar{y})^2}{\sigma^4} = \frac{1}{\sigma^4} \left[ n - \frac{3\sum_{i=1}^{n}(y_i-\bar{y})^2}{\sigma^2} \right].
\]

When \( \sigma^2 = \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{n} \),

\[
\frac{d^2\ln(l(\sigma))}{d\sigma^2} = -\frac{2n}{\sigma^2} < 0.
\]

So the m.l.e. for \( \sigma^2 \) is \( \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{n} \).