UPPSALA UNIVERSITET

Föreläsningsanteckningar

Inferensteori

Rami Abou Zahra

Inlämningsdatum
January 1, 2023

Contents

1. TODO

- Experiment in r (QQ-plot of exp vs n(0,1) data)
- Understand .dat files
- Add proof from book of theorem 4.9
- Problems 7.2.2 in the book
- Stora talens lag
- MLE better than methods of moments
- Derivatan av binomial
- Statistical significant change
- Pivotal storhet

2. Important notes from the book

2.1. Definitions/Theorems.

Definition/Sats 2.1: Mean/Medelvärde

Given n samples $x_1, \cdots, x_n$, the mean:

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

Definition/Sats 2.2: Median

Given n samples $x_1, \cdots, x_n$, the median is *the middle value* of the sorted sample

If the middle value contains 2 values (if n is even), the median is the mean of the two middle values

Definition/Sats 2.3: Mode/Typvärde

The most common number in the data set

Definition/Sats 2.4: Sample variance

Denoted by $s^2 = \sigma^2$

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Definition/Sats 2.5: Sample standard deviation/Standardavvikelse

Given by $\sqrt{s^2}$:

$$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Definition/Sats 2.6: Range/Variationsbredd

The difference between the largest number and the smallest number in the data set

Definition/Sats 2.7: Quartile/Kvartil

The median in the upper resp. lower half of the sorted data

Definition/Sats 2.8: Inter quartile range/Kvartilavstånd

The difference between the upper and lower quartile

Definition/Sats 2.9: Sample covariance/Kovarians

Let the data set be 2-dimensional tuples $(x_1, y_1), \cdots, (x_n, y_n)$:

$$c_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})$$

Anmärkning:

$$S_{xy} = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})$$

Definition/Sats 2.10: Sample correlation coefficient

Let the data set be 2-dimensional tuples $(x_1, y_1), \cdots, (x_n, y_n)$

$$r_{xy} = \frac{c_{xy}}{s_x \cdot s_y}$$

Anmärkning:

$$-1 \leq c_{xy} \leq 1$$

2.2. **Problems and Solutions.**

2.2.1. *601.*

Given $x_1, \cdots, x_5$ and $y_1, \cdots, y_9$ we have:

$$\begin{aligned} \overline{x} &= 12.2 \qquad & s_x &= 2.1 \\ \overline{y} &= 15.8 \qquad & s_y &= 2.9 \end{aligned}$$

We want to combine this data into one variable $z = x_1, \cdots, x_5, y_1, \cdots, y_9$

Calculate the mean and standard deviation for z

Solution:

The mean $\frac{1}{5+9} \sum z_i = \frac{1}{5+9} \left(\sum x_i + \sum y_i \right)$

Notice we are given the mean for each variable, through algebraic manipulation we get:

$$\overline{x} = \frac{1}{5} \sum_{i=1}^{5} x_i \Leftrightarrow 5\overline{x} = \sum_{i=1}^{5} x_i = 61$$

$$\overline{y} = \frac{1}{9} \sum_{i=1}^{9} y_i \Leftrightarrow 9\overline{y} = \sum_{i=1}^{9} y_i = 142.2$$

Therefore:

$$\overline{z} = \frac{1}{14}(61 + 142.2) = 14.514$$

The standard deviation is a little trickier, but still follows from algebraic manipulation:

$$
\begin{aligned}
s_x = 2.1 \Rightarrow s_x^2 = 4.41 = \frac{1}{5-1}\sum_{i=1}^{5}(x_i - 12.2)^2 \\
4.41 \cdot 4 = \sum_{i=1}^{5}(x_i - 12.2)^2 = \sum_{I=1}^{5}(x_i^2 - 2 \cdot 12.2 \cdot x_i + 12.2^2) \\
\Rightarrow \sum_{i=1}^{5} x_i^2 - 2 \cdot 12.2 \sum_{i=1}^{5} x_i + \sum_{i=1}^{5} 12.2^2 \\
= \sum_{i=1}^{5} x_i^2 - 2 \cdot 12.2 \underbrace{\sum_{i=1}^{5} x_i}_{5\overline{x}} + 5 \cdot 12.2^2 \\
= \sum_{i=1}^{5} x_i^2 - 12.2^2 = 4 \cdot 4.41 = 17.64 \\
\Leftrightarrow \sum_{i=1}^{5} x_i^2 = 17.64 + 5 \cdot 12.2^2 = 761.84
\end{aligned}
$$

Same done for y gives:

$$\Leftrightarrow \sum_{i=1}^{9} y_i^2 = 67.28 + 9 \cdot 15.6^2 = 2314.04$$

The variance for z:

$$
\begin{aligned}
s_z = \frac{1}{14-1}\sum_{i=1}^{14}(z_i - \overline{z})^2 = \frac{1}{13}\sum_{i=1}^{14} z_i^2 - 2 \cdot 14\overline{z}^2 + 14\overline{z^2} \\
= \frac{1}{13}\sum_{i=1}^{14} z_i^2 - 14\overline{z^2} = 3.14176
\end{aligned}
$$

2.2.2. *602.*

We essentially proceed the same way as the did for the previous problem, but take into account that we need to *remove* 19 and *add* 91.

We are given $n = 100$:

$$
\begin{aligned}
\overline{x} = 91.28 = \frac{1}{100}\sum_{i=1}^{100} x_i \\
s = 7.5
\end{aligned}
$$

Find the correct mean and standard deviation

To find the mean, we proceed as follows:

$$
\begin{aligned}
\frac{1}{100}\sum x_i = 91.28 \Leftrightarrow 91.28 \cdot 100 = \sum x_i \\
\sum x_i - 19 + 91 = 9200 \Leftrightarrow \overline{x} = \frac{1}{100}9200 = 92
\end{aligned}
$$

Using the same trick for the standard deviation:

$$
\begin{aligned}
s^2 = 56.25 = \frac{1}{100-1}\sum(x_i\overline{x})^2 \\
\Rightarrow 5568.75 = \sum(x_i^2 - 2\overline{x}x_i + \overline{x}^2) \\
\sum x_i^2 - 100\overline{x}^2 = 5568.75 \Leftrightarrow \sum x_i^2 = 5568.75 + 100\overline{x}^2 \\
= 5568.75 + 100 \cdot (91.28)^2 = 838772.59 = \sum x_i^2
\end{aligned}
$$

Here, we correct with the squares of 19 and 91 respectively, since the summands are squared

$$8838772.59 - 19^2 + 91^2 = 846692.59 = \sum x_i^2$$

Now we can start using the real values:

$$\sum x_i^2 - 100 \cdot 92^2 = 292.59$$

$$s_x = \sqrt{\frac{1}{100} 292.59} = 1.71791455$$

2.2.3. *605.*

Here things get a little trickier. It is greatly encouraged to look at example 6.10 in the book.

We begin by splitting the data into intervala 0-4, 4-8,$\cdots$ and finding the middle point of those intervals (class middle):

2	6	10	14	18	22	26	30

Looking at our data, we convert it into *how many* components are breaking in an interval, and not how many we have left (frequency):

3	7	6	4	2	1	1	1

Now we can use the estimate that $\sum x_i \approx$ the sum of the frequency (f_i)·class middle (k_i):

$$\sum_{i=1}^{8} f_i k_i = 278 \approx \sum_{i=1}^{25} x_i$$

$$\overline{x} \approx \frac{278}{25} = 11.12$$

In order to calculate the standard deviation, we need to find the variance and in order to find the variance, we need to find $\sum x_i^2$, so let us do that

That sum is the same as squaring the class middle, we therefore have:

$$\sum_{i=1}^{25} x_i^2 \approx \sum_{i=1}^{8} k_i^2 f_i = 2^6 \cdot 3 + \cdots 30^2 \cdot 1 = 4356$$

$$s_x = \sqrt{\frac{1}{25-1} \sum_{i=1}^{25} x_i^2 - 25\overline{x}^2} \approx 7.259$$

3. Important notes from the book

3.1. Definitions/Theorems.

Definition/Sats 3.1: Sample/Stickprov

A sample $x_1, \cdots, x_n$ is of size n and is an observation from the random variable $X = X_,1 \cdots, X_n$ with distribution F

Definition/Sats 3.2: Random Sample

If the random variables $X_1, \cdots, X_n$ are independent, then the sample is a *random sample*

Definition/Sats 3.3: Estimate/Skattning

Given a sample from random variables with known distribution function but unknown ”distribution function input”, an *estimate* $\theta^*(x)$ is a function of the sample attempting to decote the unknown input (parameter)

Anmärkning:
The correct value one attempts to find is denoted by θ

Definition/Sats 3.4: Estimator

The estimation observed in the previous theorem, is an observation from the *estimator*; an observation of observed values of a random variable. The estimator is what the estimate observates, denoted by $\theta^*(X)$

Definition/Sats 3.5: Bias

$$E(\theta^*(X)) - \theta$$

Definition/Sats 3.6: Random error

$$\theta^* - E(\theta^*(X))$$

Definition/Sats 3.7: Total error

$$\theta^* - \theta = E(\theta^*(X)) - \theta + \theta^* - E(\theta^*(X))$$

Definition/Sats 3.8: Unbiased/Väntevärdesriktig

$$E(\theta^*(X)) - \theta = 0$$

Definition/Sats 3.9: Efficiency

Suppose θ_1^* and θ_2^* are unbiased estimates of θ and

$$V(\theta_1^*(X)) \leq V(\theta_2^*(X))$$

Then θ_1^* is *more efficient* than θ_2^*

Definition/Sats 3.10: Standard error/Medelfel

Estimate of the standard deviation, which is $\sqrt{Var}$:

$$D(\theta^*(X)) = d(\theta^*)$$

Definition/Sats 3.11: Mean squared error

$$M(\theta^*) = E((\theta^*(X) - \theta)^2)$$

Anmärkning:
Recall that $V(X) = E(X^2) - (E(X))^2 \Leftrightarrow E(X^2) = V(X) + (E(X))^2$
Therefore, MSE can be written as $E((\theta^*(X) - \theta)^2) = V(\theta^*(X) - \theta) + \underbrace{(E(\theta^*(X) - \theta))^2}_{\text{bias}^2}$

Definition/Sats 3.12: Asymptotically unbiased/Asymptotiskt Väntevärdesriktig

If the bias $B(\theta_n^*)$ goes to 0 as $n \to \infty$, then it is *asymptotically unbiased*

(for all θ in the parameter-space)

Definition/Sats 3.13: Convergence/Konvergens

The estimator $\theta_n^*(X)$ *converges* to θ:

- **In probability**:
  - If for every $\varepsilon > 0$ $P(|\theta_n^*(X) - \theta| > 0) = 0$ as $n \to \infty$
  - Notice the comparison sign, we are saying "the probability that our estimate is off from the true value by a lot goes to zero"
- **In square means**:
  - If the mean squared error $M(\theta_n^*) \to 0$ as $n \to \infty$

Anmärkning:
If the estimator converges in square means, then it converges in probability

Definition/Sats 3.14: Consistent

The estimate θ_n^* is said to be *consistent* if the estimator $\theta_n^*(X)$ converges in probability for all θ

Definition/Sats 3.15

If the estimate θ_n^* is asymptotically unbiased and $V(\theta_n^*(X)) \to 0$ as $n \to \infty$ for all θ, then our estimate is consistent

Bevis 3.1

The mean square error goes to zero, by an earlier remark it therefore converges in probability and can be written in the following way:

$$M(\theta_n^*) = V(\theta_n^*) + B^2(\theta_n^*)$$

Since the estimate is unbiased, the bias = 0, and per the theorem, the variance goes to 0 as $n \to \infty$. Then, by theorem 6.13 it converges in square means and by the remark, it also converges in probability. $\square$

3.2. **Problems and Solutions.**

3.2.1. *7.2.1.*

We are observing events which have the same probability of happening. This is therefore a binomially distributed observation.

Since they have the same probability, one can estimate the probability as $\dfrac{\text{successfull attempts}}{\text{total attempts}} = \dfrac{4}{10} = 0.4$

In order to determine the expected value of our estimator, we look at the expected value for *any* binomially distributed chain of events:

$$E(X) = \mu = np$$

Since we are estimating $p^* = \dfrac{X}{n}$, we get the following for our expected value of our estimator:

$$E(p^*) = \frac{E(X)}{n} = p$$

4. Important notes from the book

4.1. Definitions/Theorems.

Definition/Sats 4.1: Method of moments/Momentmetoden

Let $x_1, \cdots, x_n$ random sample from X with $E(X) = m(\theta)$, where m is some known function of the unknown parameter θ
If θ is one dimensional, the moment estimate $\theta = \theta^*$ solves equation $m(\theta) = \overline{x}$

Definition/Sats 4.2

Let $x_1, \cdots, x_n$ random sample from X with $E(X) = \theta$
The estimate $\theta^* = \overline{x}$ is unbiased and if $\sigma^2 = V(X) < \infty$ then it is consistent as well

Bevis 4.1

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{n}{n}E(X_i) = \theta$$
$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = n\frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

By theorem 6.15, the estimate is unbiased (per def. in this case) and the variance goes to 0 as n increases, therefore it is consistent. $\square$

Definition/Sats 4.3: Multivariate method of moments

$\theta = (\theta_1, \theta_2)$, moment estimates solve the system:

$$E(X) = m_1(\theta_1, \theta_2) = \overline{x}$$
$$E(X^2) = V(X) + (E(X))^2 = m_2(\theta_1, \theta_2) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

Anmärkning:
If the expected value μ is known and σ^2 is the only parameter we want to estimate, then

$$(\sigma^2)^* = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Is a more efficient estimate of σ^2 rather than s^2. It is therefore unbiased and (since it is more efficient) has less variance than s^2

Definition/Sats 4.4: Sample variance is unbiased

Let $x_1, \cdots, x_n$ random sample from a random variable X with variance σ^2

The sample variance s^2 is an unbiased estimation of σ^2

Bevis 4.2: Sample variance is unbiased

Let $\mu = E(X)$. Through some algebraic manipulation, we obtain the following:

$$(x_i - \overline{x})^2 = (x_i - \mu + \mu - \overline{x})^2 = ((x_i - \mu) - (\overline{x} - \mu))^2$$
$$\Rightarrow (x_i - \mu)^2 - 2(x_i - \mu)(\overline{x} - \mu) + (\overline{x} - \mu)^2$$

Then we have S_{xx}:

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}(x_i - \mu)^2 + (\overline{x} - \mu)^2 - 2(x_i - \mu)(\overline{x} - \mu)$$
$$= \sum_{i=1}^{n}(x_i - \mu)^2 + n(\overline{x} - \mu)^2 - 2(\overline{x} - \mu)\sum_{i=1}^{n}(x_i - \mu)$$

We use (and abuse) the fact that $\overline{x} = \frac{1}{n}\sum x_i \Leftrightarrow n\overline{x}\sum x_i$:

$$\sum_{i=1}^{n}(x_i - \mu)^2 + n(\overline{x} - \mu)^2 - 2(\overline{x} - \mu)(n\overline{x} - n\mu)$$
$$= \sum_{i=1}^{n}(x_i - \mu)^2 + n(\overline{x} - \mu)^2 - 2n(\overline{x} - \mu)(\overline{x} - \mu)$$
$$\Rightarrow \sum_{i=1}^{n}(x_i - \mu)^2 - n(\overline{x} - \mu)$$

We will now look at the definition of the variance:

$$Var(X) = E((X - \mu)^2)$$

Where $\mu = E(X)$, as previous. Per assumption, $Var(X) = \sigma^2$. We have:

$$E((X_i - \mu)^2) = V(X_i) = \sigma^2$$
$$E((\overline{X} - \mu)^2) = Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \frac{1}{n^2}Var\left(\sum_{i=1}^{n}X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}Var(X_i) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

We want to show that the sample variance $s^2 = \frac{1}{n-1}S_{xx}$ is an unbiased estimation of σ^2, this equates to showing:

$$E(s^2(X)) = \sigma^2$$

Which we can show through the following:

$$E(s^2(X)) = E\left(\frac{1}{n-1}S_{xx}\right) = \frac{1}{n-1}E(S_{xx})$$
$$E(S_{xx}) = E\left(\sum_{i=1}^{n}(X_i - \mu)^2 - n(\overline{X} - \mu)^2\right)$$
$$= E\left(\sum_{i=1}^{n}(X_i - \mu)^2\right) - nE\left((\overline{X} - \mu)^2\right) = n\sigma^2 - n\frac{\sigma^2}{n} = \sigma^2(n-1)$$
$$\Rightarrow \frac{1}{n-1}E(S_{xx}) = \frac{1}{n-1}\sigma^2(n-1) = \sigma^2$$

Since $E(s^2(X)) = \sigma^2$, the estimate is unbiased.

□

Anmärkning:
If $E(X^4) < \infty$ then s^2 is a consistent estimate of σ^2

4.2. **Problems and Solutions.**

4.2.1. *7.2.3.*

Using the method of moments to estimate p^*, we have that $E(X) = m(p) = np$

Since our parameter p is one-dimensional, we have that:

$$m(p) = \overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}(1+1+1+1+0+0\cdots+0) = 0.4$$

4.2.2. *7.2.4.*

Proceeding as with the previous problem, since we have a one-dimensional parameter, we simply look at the mean:

$$\overline{x} = \frac{1}{20}(1+0\cdots+0) = \frac{1}{20} = 0.05$$

4.2.3. *7.2.5.*

This is a multivariate method of moments, i.e we would like to find $m(n,p)$. What is a little tricky about this question is *not* using Theorem 8.3, but constructing a little "diy" system of equations.

We know the random variable is binomially distributed, therefore:

$$m_1(n,p) = n \cdot p = \overline{x} = \frac{1}{2}(3+5) = 4$$

$$m_2(n,p) = n \cdot p \cdot q = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{2-1}((3-4)^2 + (5-4)^2) = 2$$

$$\begin{cases} n \cdot p = 4 \\ n \cdot p \cdot q = n \cdot p \cdot (1-p) = 2 \end{cases} \Rightarrow \begin{cases} n = 8 \\ p = 0.5 \end{cases}$$

4.2.4. *7.2.6.*

We have that $Var(s) = E(s^2) - (E(s))^2$. Notice that s^2 is an unbiased estimate for σ^2, this means that $E(s^2) = \sigma^2$.
We therefore have the following:

$$Var(s) = \sigma^2 - (E(s))^2$$
$$\Leftrightarrow (E(s))^2 = \sigma^2 - Var(s) \Rightarrow (E(s))^2 \leq \sigma^2$$
$$\Rightarrow E(s) \leq \sigma$$

In order to obtain equality, we need $Var(s) = 0$, this happens when $\exists a\ P(s = a) = 1$

5. Important notes from the book

5.1. Definitions/Theorems.

Definition/Sats 5.1: Likelihood-function

Let $x_1, \cdots, x_n$ be a random sample from the random variable X with distribution $F(x;\theta)$.
Likelihood function is defined as follows:

$$L(\theta) = \begin{cases} \prod_{i=1}^{n} p(x_i;\theta) & \text{discrete} \\ \prod_{i=1}^{n} f(x_i;\theta) & \text{continous} \end{cases}$$

Definition/Sats 5.2: Loglikelihood

Defined as follows:

$$l(\theta) = \ln L(\theta)$$

Definition/Sats 5.3: MLE

Let $x_1, \cdots, x_n$ be a random sample from the random variable X with distribution $F(x;\theta)$.

The *maximum-likelihood estimate* of θ is the θ which maximizes the likelihood function ($\Leftrightarrow$ Loglikelihood function)

Anmärkning:
Likelihood function depends on both $x_1, \cdots, x_n$ and θ, but for MLE we fix $x_1, \cdots, x_n$ and only study how θ affects the function.

Anmärkning:
An MLE is generally consistent

Anmärkning:
An MLE is not always unbiased, but can be corrected.

5.2. Problems and Solutions.

5.2.1. *7.2.7.*

In order to start our estimation, we need to remember that our observations come from a random variabe, which has some sort of distribution which we need to figure out.
Notice how Kalle plays *until* something happens, this hints that our distribution is $\sim$*ffg*, which is a discrete distribution with the following distribution function:

$$p(k) = p(1-p)^{k-1}$$

In our case, we are estimating p, our product becomes:

$$\prod_{i=1}^{7} p_i(x_i;\theta) = \prod_{i=1}^{7} p(1-p)^{x_i-1}$$

$$\Rightarrow p^7(1-p)^{(\sum x_i)-7} = p^7(1-p)^{(3+7+10+\cdot+4)-7} = p^7(1-p)^{42} = L(p)$$

In order to find which p maximizes this function, we differentiate:

$$L'(p) = -42 \cdot 7p^7(1-p)^{41}$$

$$L'(p) \Rightarrow p \in \{0,1\}$$

These are the trivial roots, let us therefore examine the loglikelihood function:

$$
\begin{aligned}
l(p) = \ln\left(L(p)\right) &= \ln\left(p^7\right) + \ln\left(1-p\right)^{42} \\
&= 7\ln(p) + 42\ln(1-p) \\
l'(p) &= \frac{7}{p} - \frac{42}{1-p}
\end{aligned}
$$

We maximize the loglikelihood:

$$
\begin{aligned}
l'(p) = 0 \Leftrightarrow \frac{7}{p} &= \frac{42}{1-p} \Rightarrow 6p = 1-p \\
p &= \frac{1}{7}
\end{aligned}
$$

5.2.2. *7.2.8.*

We are given values in the range $0 \leq x_i \leq 1$, and we therefore really only need to look at that specific case in our cumulative distribution function.
Since the distribution function in that interval is continous, it is safe to assume our random variable is continous as wel.
For a likelihood function for a continous variable, we need the probability density function, which we can obtain through taking F'_X:

$$
\begin{aligned}
F_X(x) &= x^\alpha \\
F'_X &= \alpha x^{\alpha-1}
\end{aligned}
$$

This gives the following likelihood function:

$$
\begin{aligned}
L(\alpha) = \prod_{i=1}^{10} f_x(x_i;\alpha) &= \prod_{i=1}^{10} \alpha x_i^{\alpha-1} = a^{10}\prod_{i=1}^{10} x_i^{\alpha-1} \\
&= \alpha^{10}\left(\prod_{i=1}^{10} x_i\right)^{\alpha-1}
\end{aligned}
$$

Differentiating with respect to α yields:

$$
\begin{aligned}
L'(\alpha) &= 10\alpha^9(\Pi x_i)^{\alpha-1} + a^{10}(\Pi x_i)^{\alpha-1}\ln(\Pi x_i) \\
L'(\alpha) = 0 \Leftrightarrow 10\alpha^9(\Pi x_i)^{\alpha-1} &= -a^{10}(\Pi x_i)^{\alpha-1}\ln(\Pi x_i) \\
10 &= -a\ln(\Pi x_i)
\end{aligned}
$$

We calculate $\Pi x_i = 0.57 \cdot 0.81 \cdots 0.99 \approx 0.006189$ and solve for α:

$$
\alpha = \frac{-10}{\ln(0.006189)} \Rightarrow \alpha \approx 1.966
$$

5.2.3. *7.2.9.*

Since we are given the density function, all we need to do is determine the likelihood function. We are given 9 observations:

$$
\begin{aligned}
&\prod_{i=1}^{9} f_X(x_i;\theta) = \prod_{i=1}^{9} \frac{x_i}{\theta^2} e^{-x_i/\theta} \\
\Rightarrow &\left(\frac{1}{\theta^2}\right)^9 \prod_{i=1}^{9} x_i e^{-x_i/\theta} = \frac{1}{\theta^{18}}\prod_{i=1}^{9} x_i \prod_{i=1}^{9} e^{-x_i/\theta} \\
&= \frac{1}{\theta^{18}}\left(\Pi x_i\right) e^{-(\sum x_i)/\theta} = L(\theta)
\end{aligned}
$$

Differentiate and fix maximum:

$$\begin{aligned}
L'(\theta) &= -18\theta^{-19}\left(\Pi x_i\right)e^{-(\sum x_i)/\theta} + \theta^{-18}\left(\Pi x_i\right)e^{-(\sum x_i)/\theta}\left(\frac{\sum x_i}{\theta^2}\right)\\
&= -18\theta^{-19}\left(\Pi x_i\right)e^{-(\sum x_i)/\theta} + \theta^{-20}\left(\Pi x_i\right)e^{-(\sum x_i)/\theta}\left(\sum x_i\right)\\
L'(\theta) = 0 &\Leftrightarrow \theta^{-20}\left(\Pi x_i\right)e^{-(\sum x_i)/\theta}\left(\sum x_i\right) = 18\theta^{-19}\left(\Pi x_i\right)e^{-(\sum x_i)/\theta}\\
&\Rightarrow \theta^{-20}\left(\sum x_i\right) = 18\theta^{-19}\\
&\Rightarrow \theta^{-20}\left(\sum x_i\right) = 18\theta^{-19}\\
&\Rightarrow \theta = \frac{\sum x_i}{18}
\end{aligned}$$

We calculate $\sum x_i = 23.7$ and insert:

$$\theta = \frac{23.7}{18} \approx 1.31666$$

6. Lesson 1

6.1. **727.**

Kalle lägger patiens, en gång per kväll, tills den går ut för första gången.
Under en vecka får han observationerna

$$3\ 7\ 10\ 5\ 12\ 8\ 4$$

Bestäm ML-skattningen av $p = P(\text{patiensen går ut})$

Lösning:
Här är slumpvariabeln ffg fördelad.
Låt X vara antalet gånger tills patiensen går ut, då är fördelningsfunktionen:

$$p_X(k) = (1-p)^{k-1}$$

Vi räknar med ML-skattning, vilket är:

$$\begin{aligned} L(p) &= \prod_{i=1}^{n} p_X(x_i) = \prod_{i=1}^{n} (1-p)^{x_i - 1} p \\ &= (1-p)^{\sum_{i=1}^{n} x_i - n} p^n \end{aligned}$$

Vi logaritmerar:

$$\begin{aligned} l(p) = \ln\{L(p)\} &= (\sum_{i=1}^{n} x_i - n)\ln(1-p) + n\ln(p) \\ l'(p) &= -\left(\sum_{i=1}^{n} x_i - n\right)\frac{1}{1-p} + \frac{n}{p} \\ l''(p) &= -\left(\sum_{i=1}^{n} x_i - n\right)\left(\frac{1}{1-p}\right)^2 - \frac{n}{p^2} < 0 \Rightarrow \text{ max} \\ 0 = l'(p) \Rightarrow p &= \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x}} = \frac{1}{7} \end{aligned}$$

ML-skattningen är $p^* = \dfrac{1}{7}$ vilket är samma som momentskattningen (så blir det ofta men inte alltid)

6.2. **7.210.**

En enarmad bandit (spel) med vinstchans p
Albert has spelat 10 ggr och fått 2 vinster
Beata spelade tills första vinst, gång 7
ML-skatta p

Lösning:
Låt X_1 = antal vinster under 10 spel. Denna slumpvariabel är Binomialfördelad med 10,p
Låt X_2 = antalet spel till första vinst. Denna slumpvariabel är ffg fördelad med parameter p

$$\begin{aligned}
L(p) &= p_{X_1}(x_1;p)p_{X_2}(x_2;p) \qquad x_1 = 2 \quad x_2 = 7 \\
&= \binom{10}{x_1} p^{x_i}(1-p)^{10-x_1} \cdot (1-p)^{x_2-1}p \\
&= \binom{10}{x_1} p^{x_1+1}(1-p)^{9-x_1+x_2}
\end{aligned}$$

$$l(p) = \ln\{L(p)\} = \ln\binom{10}{x_1} + (x_1+1)\ln(p) + (9-x_1+x_2)\ln(1-p)$$

$$l'(p) = (x_1+1)\frac{1}{p} - (9-x_1+x_2)\frac{1}{1-p}$$

$$l''(p) = -(x_1+1)\frac{1}{p^2} - (9-x_1+x_2)\frac{1}{(1-p)^2} < 0 \quad \text{om } 9-x_1+x_2>0 \quad \text{vilket vi har eftersom} 9-2+7>0$$

$$0 = l'(p) \Rightarrow p = \frac{x_1+1}{x_2+10} = \frac{2+1}{7+10} = \frac{3}{17} \approx 0.176$$

ML-skattningen är då $p^* = \dfrac{3}{17}$

6.3. **7.2.12.**

Taxi problemet. 7 taxibilar observeras. De är numrerade $1, \cdots, N$
Obs, numren $07o, 234, 166, 7, 65, 17, 4$

ML-skatta N

$X =$ numret på en taxibil, diskret likformigt fördelad på $(1, 2, \cdots, N)$

Sannolikhetsfunktionen är då:

$$p_X(k) = \begin{cases} \dfrac{1}{N}, 1 \leq k \leq N \\ 0, \text{ annars} \end{cases}$$

N kan vara hur stort som helst, $N \in \mathbb{N}^+ =$ rummet av alla positiva heltal

ML-skatta N (observationer $x_1, \cdots, x_n$):

$$L(N) = \prod_{i=1}^{n} p_X(x_i) = \begin{cases} \left(\dfrac{1}{N}\right)^N & \text{om } \forall x_i \leq N \\ 0 & \text{annars} \end{cases}$$

(Logaritmera/derivera funkar ej här, man måste tyvärr tänka)

ML-skattningen N^* inträffar i $\max x_i = 234$

Momentskattning: $m(n) = E(X) = \dfrac{N+1}{2}$
Lös $\overline{x} = \dfrac{N+1}{2} \Rightarrow N = 2\overline{x} - 1$
Vi har $\overline{x} = 84.3$, momentskattningen blir 167.6, vilket blir en orimlig skattning eftersom vi har en observation som är större.

6.4. **7.2.14.**

x_1, x_2 mätningar av en storhet med värdet μ
x_3 mätning av en storhet med värdet 2μ
Mätningar saknar systematiska fel, men har en slumpfel standardavvikelse σ

Bestäm MK-skattningen av μ och visa att den är väntevärdesriktigt

Minimera $Q(\mu) = (x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - 2\mu)^2$, detta kan vi lösa med derivering:

$$\begin{aligned}
Q'(\mu) &= -2(x_1 - \mu) - 2(x_2 - \mu) - 4(x_3 - 2\mu) \\
&= -2x_1 - 2x_2 - 4x_3 + 12\mu \\
Q''(\mu) &= 12 > 0 \quad \text{ger min} \\
0 = Q'(\mu) \Leftrightarrow \mu &= \frac{1}{12}(2x_1 + 2x_2 + 4x_3) = \frac{1}{6}x_1 + \frac{1}{6}x_2 + \frac{1}{3}x_3
\end{aligned}$$

MK-skattningen μ^*
Estimatorn $\mu^*(X_1 + X_2 + X_3) = \frac{1}{6}X_1 + \frac{1}{6}X_2 + \frac{1}{3}X_3$
Då blir:

$$\begin{aligned}
E\{\mu^*(X_1; X_2; X_3)\} &= \frac{1}{6}E(X_1) + \frac{1}{6}E(X_2) + \frac{1}{3}E(X_3) \\
&= \frac{1}{6}\mu + \frac{1}{6}\mu + \frac{1}{3}2\mu = \mu
\end{aligned}$$

Då är μ^* väntevärdesriktigt

En annan skattning är:

$$\mu' = \frac{2x_1 + 2x_2 + x_3}{6}$$

Är den väntevärdesriktigt? Vi tittar på motsvarande estimator:

$$\begin{aligned}
\mu'(X_1, X_2, X_3) &= \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{6}X_3 \\
E\{\mu'(X_1, X_2, X_3)\} &= \frac{1}{3}E(X_1) + \frac{1}{3}E(X_2) + \frac{1}{6}E(X_3) \\
&= \frac{1}{3}\mu + \frac{1}{3}\mu + \frac{1}{6}2\mu = \mu
\end{aligned}$$

Ok!

Vilken skattning är effektivast?
Då jämför vi variansena:

$$\begin{aligned}
V(\mu) &= \frac{6}{36}\sigma^2 \\
V(\mu') &= \frac{9}{36}\sigma^2
\end{aligned}$$

6.5. **702.**

Observationer: 4.0, 1.1, 0.2, 1.2, 2.5, 2.0, 0.7, 1.0 är ett stickprov från en Raylerghfördelning med täthetsfunktion:

$$F_X(x) = axe^{-\dfrac{ax^2}{2}} \qquad x \leq 0$$

ML-skatta a:

$$L(a) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} a x_i e^{-\dfrac{a x_i^2}{2}}$$

$$= a^n \left(\prod_{i=1}^{n} x_i \right) e^{-\dfrac{a}{2} \sum_{i=1}^{n} x_i^2}$$

$$l(a) = \ln\{L(a)\} = n \ln(a) + \sum_{i=1}^{n} \ln x_i - \frac{a}{2} \sum_{i=1}^{n} x_i^2$$

$$l'(a) = \frac{n}{a} - \frac{1}{2} \sum_{i=1}^{n} x_i^2$$

$$l''(a) = -\frac{n}{a^2} < 0 \Rightarrow \ \max$$

$$0 = l'(a) \Rightarrow a = \frac{2n}{\sum_{i=1}^{n} x_i^2}$$

ML skattning $a^* = \dfrac{2n}{\sum_{i=1}^{n} x_i^2} = \dfrac{2 \cdot 8}{30.43} \approx 0.526$

7. Important notes from the book

7.1. Definitions/Theorems.

Definition/Sats 7.1: Least Squares Estimate (LSE) / Minsta-kvadrat skattning

Let $x_1, \cdots, x_n$ be a random sample from X with $E(X) = m(\theta)$, where m is some known function. Let:

$$Q(\theta) = \sum_{i=1}^{n} (x_i - m(\theta))^2$$

The θ that minimizes Q is the *LSE* of θ

Anmärkning:
Suppose θ is one-dimensional and $\exists m'(\theta)$ such that $m'(\theta) \neq 0$, then:

$$Q'(\theta) = -2m'(\theta) \sum_{i=1}^{n} (x_i - m(\theta))$$

Finding minimum becomes (since $m'(\theta) \neq 0$):

$$\begin{aligned} &\sum_{i=1}^{n} (x_i - m(\theta)) = 0 \\ \Rightarrow &\left(\sum_{i=1}^{n} x_i \right) - nm(\theta) = 0 \\ \Rightarrow &\sum_{i=1}^{n} x_i = nm(\theta) \\ \Rightarrow &\,\overline{x} = m(\theta) \end{aligned}$$

LSE is given by $m(\theta) = \overline{x}$, just as in the case of method of moments.

Definition/Sats 7.2: Weighted LSE

If our random sample comes from different random variables with expected values $m_i(\theta)$ and have *different* standard deviation, then:

$$Q(\theta) = \lambda \cdot \sum_{i=1}^{n} \left(\frac{x_i - m_i(\theta)}{\sigma_i} \right)^2 \qquad \lambda \in \mathbb{R}$$

This is called the *Weighted Least Squares Estimate.*
λ is some constant.

Anmärkning:
If our random sample comes from different random variables with expected value $m_i(\theta)$ but all have the *same* standrad deviation, then:

$$Q(\theta) = \sum_{i=1}^{n} (x_i - m_i(\theta))^2$$

7.2. Problems and Solutions.

7.2.1. *7.2.13.*

We use the variance trick, namely:

$$Var(X) = E(X^2) - (E(X))^2$$

Since $Y = X^2$, we are essentially looking for:

$$E(X^2) = Var(X) + (E(X)) = \sigma^2 + \mu_X = 1 + \mu^2$$

Our LSE function becomes:

$$Q(\mu) = \sum_{i=1}^{10} (y_i - (\mu^2 + 1))^2$$

$$= \sum_{i=1}^{10} \left(y_i^2 - 2(\mu^2 + 1)y_i + (\mu^2 + 1)^2\right) = \left(\sum_{i=1}^{10} y_i^2\right) - 2(\mu^2 + 1)\left(\sum y_i\right) + 10(\mu^2 + 1)^2$$

We differentiate and find minimum of Q:

$$Q'(\mu) = 20(\mu^2 + 1)(2\mu) - 2\left(\sum y_i\right)(2\mu) = 0$$

$$\Rightarrow 40(\mu^2 + 1)\mu = 4\left(\sum y_i\right)\mu$$

$$\Leftrightarrow 10(\mu^2 + 1) = \sum y_i \Rightarrow \mu = \sqrt{\frac{\sum y_i}{10} - 1}$$

We calculate $\sum y_i = 0.17 + 0.06 + \cdots + 2.1 = 22.35$ and insert:

$$\mu = \sqrt{2.235 - 1} \approx 1.1113$$

7.2.2. *7.2.14.*

See lesson 1

7.2.3. *7.2.15.*

Find the averages of the observed data:

- $m_1 + m_2 = 44.65$
- $m_1 + m_3 = 50.01$
- $m_2 + m_3 = 54.87$

Then, use Gauss-elimination to solve the following system of equations:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 44.65 \\ 50.01 \\ 54.87 \end{bmatrix}$$

8. Important notes from the book

8.1. **Definitions/Theorems.**

Anmärkning 8.1

For normally distributed data, we have the following estimate for variance:

$$(\sigma^2)^* = \frac{S_{xx}}{n-1} = s^2$$

This is however assuming that μ is unknown and therefore estimated using $\mu = \overline{x}$.
If we are given μ, then the following estimate is unbiased and consistent:

$$(\sigma^2)^* = \frac{\sum(x_i - \mu)^2}{n}$$

Anmärkning 8.2

For normally distributed data from $X \sim N(\mu_1, \sigma^2)$ and $Y \sim N(\mu_2, \sigma^2)$, we can estimate σ^2 through (and using $\mu_1^* = \overline{x}$ and $\mu_2^* = \overline{y}$):

$$(\sigma^2)^* = \frac{\sum_{i=1}^{n_1}(x_i - \overline{x})^2 + \sum_{i=1}^{n_2}(y_i - \overline{y})^2}{n_1 + n_2 - 2}$$

This is an unbiased estimate

Anmärkning:
Anmärkning 8.2 raises the question of knowing *when* the variance would be equal.
If we are looking at 2 different observations, then the variation in those observations depend on the method of analysis chosen, and we have used the same method of analysis with unknown σ for both samples.

Anmärkning 8.3

For a sample from a hypergeometric distribution, it is very difficult to maximize the likelihood function.
Through numerical observations, we may use the following estimate:

$$p^* \approx \frac{x}{n}$$

Anmärkning 8.4

Suppose you have 2 samples from X and Y who both have the same distribution function, and same variance
We can then find the likelihood-function as a product of the likelihood-function for the two different samples:

$$L(\theta) = L_X(\theta) \cdot L_Y(\theta)$$
$$l(\theta) = l_X(\theta) + l_Y(\theta)$$

8.2. **Problems and solutions.**

8.2.1. *7.2.17.*

Using the unbiased formula for two normally distributed samples with the same variance, we have:

$$(\sigma^2)^* = \frac{\overbrace{\sum_{i=1}^{n_1}(x_i - \overline{x})^2}^{S_{XX}} + \overbrace{\sum_{i=1}^{n_2}(y_i - \overline{y})^2}^{S_{YY}}}{n_1 + n_2 - 2}$$

Recall that $S_{XX} = (n-1) \cdot s_x^2$, this gives:

$$\left.\begin{aligned} S_{XX} &= (5-1) \cdot 2.85 = 11.4 \\ S_{YY} &= (9-1) \cdot 3.525 = 28.2 \end{aligned}\right\} \Rightarrow \frac{11.4 + 28.2}{9 + 5 - 2} = 3.3$$

8.2.2. *7.2.18.*

The samples are from a binomial distribution (and same probability for each outcome). We can therefore let $z = x + y$ and $n = n_1 + n_2$. The estimate then becomes:

$$\frac{z}{n} = \frac{x+y}{n_1+n_2} = \frac{15+25}{100+200} = \approx 0.13$$

In orde to find the standard error, we wish to find $\sqrt{s^2/n^2}$ where $(\sigma^2)^* = s^2$

$$(\sigma^2)^* = 300 \cdot 0.13 \cdot (1 - 0.13) \Rightarrow \sqrt{\frac{34.66}{300^2}} \approx .0.0196$$

8.2.3. *7.2.19.*

Here we combine the number of cars (20+40=60) and the number of minutes (60+90=150) and use this:

$$\begin{aligned} \#\text{ of cars } &= \lambda \cdot t \Leftrightarrow 60 = 150 \cdot \lambda \\ &\Rightarrow \lambda = 0.4 \end{aligned}$$

The standard error is given by $\sqrt{\dfrac{\lambda}{n}}$:

$$= \sqrt{\frac{0.4}{150}} \approx 0.0516$$

8.2.4. *703.*

8.2.5. *705.*

8.2.6. *723.*

9. Important notes from the book

9.1. Definitions/Theorems.

Definition/Sats 9.1: Confidence interval

The interval $I_\theta = (\underline{\theta}(x), \overline{\theta}(x))$ is called the *confidence interval* for θ with degree $1-\alpha$ if:

$$P(\underline{\theta}(X) < \theta < \overline{\theta}(X)) = 1 - \alpha$$

This is a two-sided interval

Anmärkning:
α is sometimes referred to as the error-risk (felrisk)

Anmärkning:
Think of α like the risk of having an interval *not* containing θ.

Anmärkning:
The error-risk in a two-sided interval is distributed evenly as follows:

$$P(\underline{\theta}(X) < \theta) = P(\theta < \overline{\theta}(X)) = \frac{\alpha}{2}$$

Anmärkning:
We need to respect the values that the estimator attains in our intervals.
Therefore, it is sometimes more reasonable to have a one-sided interval (if for example the estimator is always positive)

Definition/Sats 9.2: Reference variable/Pivot variable/Referensvariabel

Usually denoted by R_θ is a variable such that:

- Only depends on θ
- Has completely known (known values) distribution

Anmärkning:
Normally when using CLT, our reference varible looks like:

$$R_\mu = \frac{\mu^* - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Definition/Sats 9.3: Construction of confidence interval

- Estimate θ using θ^*
- Find a reference variable R_θ based on the estimate $\theta^*(X)$
- Enclose R_θ using quantiles r: $P(r_{1-\alpha/2} < R_\theta < r_{\alpha/2}) = 1 - \alpha$
- Rewrite the inequality such that θ is isolated: $P(\underline{\theta}(X) < \theta < \overline{\theta}(X))$
- $I_\theta = (\underline{\theta}(X), \overline{\theta}(X))$

Definition/Sats 9.4: Confidence interval for function $\theta = g(\mu)$

Assume $I_\mu = (\underline{\mu}, \overline{\mu})$ with degree $1-\alpha$ and $\theta = g(\mu)$:

- If g is a monotely increasing function, then $I_\theta = (g(\underline{\mu}), g(\overline{\theta}))$ is a confidence interval for θ with degree $1-\alpha$

- If g is a monotely decreasing function, then $I_\theta = (g(\overline{\theta}), g(\underline{\mu}))$ is a confidence interval for θ with degree $1-\alpha$

9.2. **Problems and Solutions.**

9.2.1. *7.3.1.*

9.2.2. *7.3.2.*

9.2.3. *7.3.3.*

10. Important notes from the book

10.1. Definitions/Theorems.

Definition/Sats 10.1: Null-hypothesis

The *null-hypothesis*, denoted by H_0 is a hypothesis we make of the parameter θ, often $H_0 : \theta = \theta_0$

The *alternative hypothesis*, usually denoted by H_1 is the alternative hypothesis we are testing for. The alternative hypothesis can be *simple* ($H_1 : \theta = \theta_0$), or *composite* (eg. $H_1 : \theta > \theta_0$).

A hypothesis on the form $\theta > \theta_0$ or $\theta < \theta_0$ is called *one-sided*, while $\theta \neq \theta_0$ is a *two-sided hypothesis*

Definition/Sats 10.2: Significance level

The **significance level** α is defined as the probability to reject the null-hypothesis:

$$\alpha = P_{H_0}(\text{reject } H_0)$$

Normally $\alpha = 0.05$ or $\alpha = 0.01$ or $\alpha = 0.001$

Definition/Sats 10.3: Power of test/Styrkefunktion

The *power* of a hypothesis test is given by the power-function:

$$h(\theta) = P_\theta(\text{reject } H_0)$$

Anmärkning:
Observe that $h(\theta_0) = \alpha$ and that $h(\theta)$ is large when $\theta \in H_1$

Definition/Sats 10.4: Test variable method

Find a *testvariable* $T(x)$ based on the sample and find the critical area C for that testvariable. The test becomes:

- If $T \in C$, we reject H_0 and say ”we have a significant result”
- If $T \notin C$, we cannot reject H_0 and say ”the result is not significant”

We pick the critical area such that the significance level α is the one we want

Anmärkning:
Just because we do not reject H_0, does **not** mean we reject H_1. There just is not enough empirical data.

Definition/Sats 10.5: Direct method

The *direct method* is also based on a testvariable, but the ciritcal area is not explicitly defined. Instead, we find values of the testvariable that give the emprical evidence that H_1 is true

We then calculate the P value of the test:

$$P = P_{H_0}(\text{at least equally extreme case as the observed})$$

H_0 is rejected if $P \leq \alpha$

Anmärkning:
The direct method is equivalent to the test variable method with the advantage that it also lets us know at what significance level the null-hypothesis is rejected

Definition/Sats 10.6: Confidence method

The *confidence method* can be describe as the following:

- Find the confidence interval I_θ for the parameter with the same significance level that we want from the test (degree $1-\alpha$)
- Reject $H_0 : \theta = \theta_0$ if $\theta_0 \notin I_\theta$

Anmärkning 10.1: How to choose method

- If you already have the confidence interval (either given or calculated), then use the confidence method
- Direct method is more appropriate when the observations come from a discrete variable (but can absolutely be used in other cases since the P value gives a mesuare on the magnitude of the significance)

Anmärkning 10.2: How to choose significance level

- If rejecting H_0 poses a great risk, then it is research ethic to choose a low significance level
- If a trial is resource heavy, then a low significance level is not sustainable and or desirable (even though it is the best option)
- If not sure, let $\alpha 0.05$

Anmärkning 10.3: How to choose testvariable

- Find a reference variable R_θ with the following properties:
  - Completely known distribution
  - Only depends on θ
- Choose the testvariable $T = R_\theta$. Now T can be computed from a completely known distribution
- Choose a ciritical area C adapted after the alternative hypothesis H_1 using the known quantiles for the distribution ($C = \{T \geq r_\alpha\}$)

Anmärkning 10.4: How to choose ciritcal area

This completely depends on how you choose your alternative hypothesis H_1
The "rule of thumb" is that the have values of the testvariable that indicate that H_1 is true should be in the critical area

Anmärkning:
You need to chose your hypothesis *before* collecting your data

Anmärkning:
It may be confusing with hypothesis-testing and confidence intervals. These are in fact very similar, and we will therefore treat them as such.

Definition/Sats 10.7: Likelihood quotient test

Let $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$
Let $L(\theta)$ be the likelihood function, and let T be the testvariable where
$$T = \frac{L_1}{L_0} = \frac{L(\theta_1)}{L(\theta_0)}$$
Reject H_0 if $T \geq K$ where $K \in \mathbb{R}$ such that we achieve our chosen α

Anmärkning 10.5

Normally we define:
$$L_1 = \sup_{\theta \in H_1} \{L(\theta)\} \qquad L_0 = \sup_{\theta \in H_0} \{L(\theta)\}$$

Anmärkning:
Previously, we had a fixed $x_1, \cdots, x_n$ from a random variable X and studied how θ affects our function. With the likelihood quotient test we fix θ, and study how our sample $x_1, \cdots, x_n$ affects our function

Anmärkning:
If you are having trouble finding a suitable test-variable (maybe you are having difficulties finding a reference variable), then we can use the likelihood quotient test

10.2. Problems and Solutions.

10.2.1. *7.4.1.*

10.2.2. *7.4.2.*

10.2.3. *7.4.4.*