Outline

1. Introduction

2. Some important concepts
   - Estimation
   - Hypothesis testing

3. Example 1: binomial data

4. Example 2: normal data
Outline

1. Introduction

2. Some important concepts
   - Estimation
   - Hypothesis testing

3. Example 1: binomial data

4. Example 2: normal data
Statistical Inference

- There are three steps for Statistical methods.
  - Data collection.
  - Data presentation
  - Data analysis.

  We focus on the third and final step - the inference.

- Seek to draw conclusions based on the data.

- Important aspect - the underlying model.
There are three steps for Statistical methods.

- Data collection.
- Data presentation.
- Data analysis.

We focus on the third and final step - the inference.

- Seek to draw conclusions based on the data.
- Important aspect - the underlying model.
Statistical Inference

- There are three steps for Statistical methods.
  - Data collection.
  - Data presentation
  - Data analysis.

We focus on the third and final step - the inference.

- Seek to draw conclusions based on the data.
- Important aspect - the underlying model.
Parametric model

- **Prior belief or notion** dictates the choice of model.
- Sometimes, a **glance at the plot** shows why some specific model may be of interest.
Parametric model

- Prior belief or notion dictates the choice of model.
- Sometimes, a glance at the plot shows why some specific model may be of interest.
What is the possible underlying model?

A linear fit!!!
What is the possible underlying model?

A linear fit !!!!
A quadratic fit might be the winner?
Outline

1. Introduction

2. Some important concepts
   - Estimation
   - Hypothesis testing

3. Example 1: binomial data

4. Example 2: normal data
Quantities of interest

- **Parameter**: some unknown but fixed quantity. Does not depend on data.
- **Statistic**: A quantity that depends on data. Computed from the sample.
- **Estimator**: A statistic to predict/substitute the parameter.
Quantities of interest

- **Parameter**: some unknown but fixed quantity. Does not depend on data.
- **Statistic**: A quantity that depends on data. Computed from the sample.
- **Estimator**: A statistic to predict/substitute the parameter.
Quantities of interest

- **Parameter**: some unknown but fixed quantity. Does not depend on data.
- **Statistic**: A quantity that depends on data. Computed from the sample.
- **Estimator**: A statistic to predict/substitute the parameter.
Statistical methods

There are two main problems of statistical analysis.

- Estimation
- Testing of hypothesis.

We will briefly discuss them here. Our examples will illustrate the difference between them.
Introduction to Statistical Inference

Some important concepts

Statistical methods

- There are two main problems of statistical analysis.
  - Estimation
  - Testing of hypothesis.
- We will briefly discuss them here. Our examples will illustrate the difference between them.
Statistical methods

There are two main problems of statistical analysis.
  - Estimation
  - Testing of hypothesis.

We will briefly discuss them here. Our examples will illustrate the difference between them.
Estimation

Estimation can be of two types.

- **Point estimation**: We seek to specify a predictive value for the parameter.
- **Interval estimation**: The goal is to specify a range of candidate values for the parameter.

We try to discuss them briefly using examples.
Estimation

Estimation can be of two types.

- **Point estimation**: We seek to specify a predictive value for the parameter.
- **Interval estimation**: The goal is to specify a range of candidate values for the parameter.

We try to discuss them briefly using examples.
Point estimation

- **Consider the following example:**
  The North Carolina State University seeks to figure out the fraction of monthly expenses spent by its students on different categories.

- **Question:** What is the percentage spent on groceries and merchandize?

- **Data:** We collect data on 1000 random students across the campus and record their expenditure pattern.
Some important concepts

Point estimation

Consider the following example:
The North Carolina State University seeks to figure out the fraction of monthly expenses spent by its students on different categories.

Question: What is the percentage spent on groceries and merchandize?

Data: We collect data on 1000 random students across the campus and record their expenditure pattern.
Point estimation

Consider the following example:
The North Carolina State University seeks to figure out the fraction of monthly expenses spent by its students on different categories.

Question: What is the percentage spent on groceries and merchandize?

Data: We collect data on 1000 random students across the campus and record their expenditure pattern.
Point estimation

- Consider the following example:
  The North Carolina State University seeks to figure out the fraction of monthly expenses spent by its students on different categories.

- Question: What is the percentage spent on groceries and merchandize?

- Data: We collect data on 1000 random students across the campus and record their expenditure pattern.
Introduction to Statistical Inference

Some important concepts

Estimation

Point estimation

- We observe that the **average spent on the purchases is 21%**.
- Parameter: the unknown fraction spent on them.
  - Statistic: average of the proportions in the 1000 students.
- This average is an **estimator** of the unknown parameter.
- This is known as **point estimation**.
- However, this does not tell us about **how close we are** to the actual fraction, or how **accurate** our estimator is.
Point estimation

- We observe that the **average spent on the purchases** is **21%**.
- **Parameter**: the unknown fraction spent on them.
  - **Statistic**: average of the proportions in the 1000 students.
- This average is an **estimator** of the unknown parameter.
- This is known as **point estimation**.
- However, this does not tell us about **how close we are** to the actual fraction, or how **accurate** our estimator is.
Point estimation

- We observe that the average spent on the purchases is 21%.
- Parameter: the unknown fraction spent on them.
  Statistic: average of the proportions in the 1000 students.
- This average is an estimator of the unknown parameter.
- This is known as point estimation.
- However, this does not tell us about how close we are to the actual fraction, or how accurate our estimator is.
Point estimation

■ We observe that the **average spent on the purchases** is 21%.

■ **Parameter** : the unknown fraction spent on them. 
  **Statistic** : average of the proportions in the 1000 students.

■ This average is an **estimator** of the unknown parameter.

■ This is known as **point estimation**.

■ However, this does not tell us about **how close we are to the actual fraction**, or how **accurate** our estimator is.
We observe that the **average spent on the purchases** is 21%.  

- **Parameter**: the unknown fraction spent on them.  
  - **Statistic**: average of the proportions in the 1000 students.  
- This average is an **estimator** of the unknown parameter.  
- This is known as **point estimation**.  
- However, this does not tell us about **how close we are** to the actual fraction, or how **accurate** our estimator is.
Interval estimation

- Instead of specifying the value at a point, one looks for a range of values as plausible, e.g. 19% to 23%.
- The goal is to ascertain some probability for such an interval, or ideally find an interval with a pre-specified probability (like .95 or .99) attached to it.
Interval estimation

- Instead of specifying the value at a point, one looks for a range of values as plausible, e.g. 19% to 23%.
- The goal is to ascertain some probability for such an interval, or ideally find an interval with a pre-specified probability (like .95 or .99) attached to it.
Introduction to Statistical Inference

Hypothesis testing

Our next example comes from the D.M.V.

The D.M.V. wants to apprise the effect of air-bags in reducing the risk of death in road accidents.

Question: Does having air-bag reduce the chance of death in collisions?

Hypothesis 1: The chances remain same. This is known as the null hypothesis (status quo)

Hypothesis 2: The risk is less for cars having air-bags. This is the alternate hypothesis.
Hypothesis testing

- Our next example comes from the D.M.V. The D.M.V. wants to apprise the effect of air-bags in reducing the risk of death in road accidents.
- **Question**: Does having air-bag reduce the chance of death in collisions?
- **Hypothesis 1**: The chances remain same. This is known as the null hypothesis (status quo)
- **Hypothesis 2**: The risk is less for cars having air-bags. This is the alternate hypothesis.
Hypothesis testing

- Our next example comes from the D.M.V. The D.M.V. wants to apprise the effect of air-bags in reducing the risk of death in road accidents.

- **Question**: Does having air-bag **reduce** the chance of death in collisions?

- **Hypothesis 1**: The chances remain same. This is known as the **null hypothesis** (status quo)

- **Hypothesis 2**: The risk is less for cars having air-bags. This is the **alternate hypothesis**.
Hypothesis testing

- Our next example comes from the D.M.V. The D.M.V. wants to apprise the effect of air-bags in reducing the risk of death in road accidents.
- **Question**: Does having air-bag reduce the chance of death in collisions?
- **Hypothesis 1**: The chances remain same. This is known as the null hypothesis (status quo)
- **Hypothesis 2**: The risk is less for cars having air-bags. This is the alternate hypothesis.
The data from last year shows about 11% among the air-bag car occupants succumb to fatal injuries.

In the cars without the safety equipments, the corresponding figure is 14%.

Question: Is the rise in percentage significant to conclude in favor of Hypothesis 2? Or is this just a chance variation, and can Hypothesis 1 not be overwhelmingly ruled out?
Hypothesis testing

- The data from last year shows about 11% among the air-bag car occupants succumb to fatal injuries.
- In the cars without the safety equipments, the corresponding figure is 14%.
- Question: Is the rise in percentage significant to conclude in favor of Hypothesis 2? Or is this just a chance variation, and can Hypothesis 1 not be overwhelmingly ruled out?
Hypothesis testing

- The data from last year shows about 11% among the air-bag car occupants succumb to fatal injuries.
- In the cars without the safety equipments, the corresponding figure is 14%.
- **Question**: Is the rise in percentage **significant** to conclude in favor of Hypothesis 2? Or is this just a chance variation, and can Hypothesis 1 not be overwhelmingly ruled out?
We will discuss the estimation procedure with a few more examples.
Outline

1. Introduction

2. Some important concepts
   - Estimation
   - Hypothesis testing

3. Example 1: binomial data

4. Example 2: normal data
A study is conducted among the Duke University students. 5 undergraduates are chosen at random and asked whether they receive their spending money from part-time jobs.

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Year</th>
<th>Part-timer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesley Pickering</td>
<td>19</td>
<td>Junior</td>
<td>NO</td>
</tr>
<tr>
<td>Jason Gullian</td>
<td>18</td>
<td>Freshman</td>
<td>YES</td>
</tr>
<tr>
<td>Erin McClintic</td>
<td>20</td>
<td>Junior</td>
<td>NO</td>
</tr>
<tr>
<td>Stacey Culp</td>
<td>19</td>
<td>Sophomore</td>
<td>NO</td>
</tr>
<tr>
<td>Fred Almirall</td>
<td>21</td>
<td>Senior</td>
<td>YES</td>
</tr>
</tbody>
</table>
A study is conducted among the Duke University students. 5 undergraduates are chosen at random and asked whether they receive their spending money from part-time jobs.

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Year</th>
<th>Part-timer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesley Pickering</td>
<td>19</td>
<td>Junior</td>
<td>NO</td>
</tr>
<tr>
<td>Jason Gullian</td>
<td>18</td>
<td>Freshman</td>
<td>YES</td>
</tr>
<tr>
<td>Erin McClintic</td>
<td>20</td>
<td>Junior</td>
<td>NO</td>
</tr>
<tr>
<td>Stacey Culp</td>
<td>19</td>
<td>Sophomore</td>
<td>NO</td>
</tr>
<tr>
<td>Fred Almirall</td>
<td>21</td>
<td>Senior</td>
<td>YES</td>
</tr>
</tbody>
</table>
Example 1: binomial data

Estimation

- An unknown fraction $\pi$ of the pool of students have part-time jobs.
- We want to estimate that unknown $\pi$.
- Given any value of $\pi \in [0, 1]$, what are the chances of having a random sample of 5 students with 2 of them doing such jobs?
- Also, what is the most likely value of $\pi$ that can generate such a sample?
- Normal guess: $\pi = \frac{2}{5} = .4.$
Estimation

- An unknown fraction $\pi$ of the pool of students have part-time jobs.
- We want to estimate that unknown $\pi$.
- Given any value of $\pi \in [0, 1]$, what are the chances of having a random sample of 5 students with 2 of them doing such jobs?
- Also, what is the most likely value of $\pi$ that can generate such a sample?
- Normal guess: $\pi = \frac{2}{5} = .4$. 

Estimation

- An unknown fraction $\pi$ of the pool of students have part-time jobs.
- We want to estimate that unknown $\pi$.
- Given any value of $\pi \in [0, 1]$, what are the chances of having a random sample of 5 students with 2 of them doing such jobs?
- Also, what is the most likely value of $\pi$ that can generate such a sample?
- Normal guess: $\pi = \frac{2}{5} = .4.$
This comes from a binomial distribution, which is discrete in nature.

If we know the value of \( \pi \), we can write the probability of the sequences.

If \( \pi = .2 \), the probability of such an occurrence is 
\[
.8 \times .2 \times .8 \times .8 \times .2 = .02048.
\]

If \( \pi = .5 \), the probability becomes 
\[
.5 \times .5 \times .5 \times .5 \times .5 = .03125.
\]

Conclusion: \( \pi = .5 \) is more likely to \( \pi = .2 \).
What is more likely?

- This comes from a **binomial distribution**, which is **discrete** in nature.

- If we know the value of \( \pi \), we can write the **probability** of the sequences.

- If \( \pi = .2 \), the probability of such an occurrence is 
  \[
  .8 \times .2 \times .8 \times .8 \times .2 = .02048.
  \]

- If \( \pi = .5 \), the probability becomes
  \[
  .5 \times .5 \times .5 \times .5 \times .5 = .03125.
  \]

- **Conclusion**: \( \pi = .5 \) is more likely to \( \pi = .2 \).
What is more likely?

- This comes from a binomial distribution, which is discrete in nature.
- If we know the value of $\pi$, we can write the probability of the sequences.
- If $\pi = .2$, the probability of such an occurrence is $\frac{1}{2} \times .2 \times \frac{1}{2} \times .8 \times .2 = .02048$.
- If $\pi = .5$, the probability becomes $\frac{1}{2} \times .5 \times \frac{1}{2} \times .5 \times .5 = .03125$.
- Conclusion: $\pi = .5$ is more likely to $\pi = .2$. 
Maximum likelihood estimator

- Of all possible values of $\pi \in [0, 1]$, which one has the largest possibility of producing the data?
- In particular, which value of $\pi$ has the largest likelihood?
- It will be called the maximum likelihood estimator.
Maximum likelihood estimator

- Of all possible values of $\pi \in [0, 1]$, which one has the largest possibility of producing the data?
- In particular, which value of $\pi$ has the largest likelihood?
- It will be called the maximum likelihood estimator.
Maximum likelihood estimator

- Of all possible values of $\pi \in [0, 1]$, which one has the largest possibility of producing the data?
- In particular, which value of $\pi$ has the largest likelihood?
- It will be called the maximum likelihood estimator.
Example 1: binomial data

For some specific value of $\pi$, the probability of the data is given by

$$L(\pi) = (1 - \pi)\pi(1 - \pi)(1 - \pi)^2$$

$$= \pi^2 - 3\pi^3 + 3\pi^4 - \pi^5$$

$$\Rightarrow \frac{d}{d\pi} L(\pi) = 2\pi - 9\pi^2 + 12\pi^3 - 5\pi^4$$

$$= \pi(1 - \pi)^2(2 - 5\pi)$$
MLE computation

- Therefore, \( \frac{d}{d\pi} L(\pi) = 0 \) if \( \pi = 0, 1 \) or \( \frac{2}{5} \).
- Now, \( L(0) = 0 = L(1) \).
- Further,
  \[
  \frac{d^2}{d\pi^2} L(\pi) = 2 - 18\pi + 36\pi^2 - 20\pi^3 = 2(1 - \pi)(1 - 8\pi + 10\pi^2)
  \]
  which is negative if \( \pi = .4 \).
- Therefore, \( \pi = .4 \) is the MLE. We denote it by \( \hat{\pi} \).
MLE computation

- Therefore, \( \frac{d}{d\pi} L(\pi) = 0 \) if \( \pi = 0, 1 \) or \( \frac{2}{5} \).
- Now, \( L(0) = 0 = L(1) \).
- Further,

\[
\frac{d^2}{d\pi^2} L(\pi) = 2 - 18\pi + 36\pi^2 - 20\pi^3 = 2(1 - \pi)(1 - 8\pi + 10\pi^2)
\]

which is negative if \( \pi = 0.4 \).
- Therefore, \( \pi = 0.4 \) is the MLE. We denote it by \( \hat{\pi} \).
MLE computation

- Therefore, \( \frac{d}{d\pi} L(\pi) = 0 \) if \( \pi = 0, 1 \) or \( \frac{2}{5} \).
- Now, \( L(0) = 0 = L(1) \).
- Further, \[
\frac{d^2}{d\pi^2} L(\pi) = 2 - 18\pi + 36\pi^2 - 20\pi^3 = 2(1 - \pi)(1 - 8\pi + 10\pi^2)
\]
  which is negative if \( \pi = 0.4 \).
- Therefore, \( \pi = 0.4 \) is the MLE. We denote it by \( \hat{\pi} \).
How likely are other plausible values of $\pi$

- We will plot $L(\pi)$ against $\pi$.
- MATLAB code:

  ```matlab
  p = [0:.01:1];
  L = (p.^2).*((1-p).^3);
  plot(p,L);
  xlabel('\pi');
  ylabel('Likelihood $L(\pi)$');
  ``

  The commands have been saved in example1.m.
Plot of $L(\pi)$
Exercise 1

- You have a bag of marbles, millions in number.
- A fraction $\pi$ of them are white, rest are red.
- You draw 10 at random, and the colors turn out to be W.R.W.W.W.R.R.W.R.W.
- Compute the MLE for $\pi$ and use MATLAB to see how likely are the other values between $[0, 1]$. 
Exercise 1

- You have a bag of marbles, millions in number.
- A fraction $\pi$ of them are white, rest are red.
- You draw 10 at random, and the colors turn out to be W.R.W.W.R.R.W.R.W.
- Compute the MLE for $\pi$ and use MATLAB to see how likely are the other values between [0, 1].
Exercise 1

- You have a bag of marbles, millions in number.
- A fraction $\pi$ of them are white, rest are red.
- You draw 10 at random, and the colors turn out to be W.R.W.W.W.R.R.W.R.W.
- Compute the MLE for $\pi$ and use MATLAB to see how likely are the other values between [0, 1].
Outline

1. Introduction

2. Some important concepts
   - Estimation
   - Hypothesis testing

3. Example 1: binomial data

4. Example 2: normal data
**Guyana rainfall data**

- We have data that can be modeled as coming from a Normal distribution with mean $\mu$ and standard deviation $\sigma$ (both unknown).
- In Guyana, South America, we record the annual rainfall in the last 6 years.

<table>
<thead>
<tr>
<th>Year</th>
<th>rainfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>95&quot;</td>
</tr>
<tr>
<td>2002</td>
<td>118&quot;</td>
</tr>
<tr>
<td>2003</td>
<td>85&quot;</td>
</tr>
<tr>
<td>2004</td>
<td>154&quot;</td>
</tr>
<tr>
<td>2005</td>
<td>102&quot;</td>
</tr>
<tr>
<td>2006</td>
<td>96&quot;</td>
</tr>
</tbody>
</table>
Guyana rainfall data

- We have data that can be modeled as coming from a **Normal distribution** with mean $\mu$ and standard deviation $\sigma$ (both unknown).

- In **Guyana, South America**, we record the **annual rainfall in the last 6 years**.

<table>
<thead>
<tr>
<th>Year</th>
<th>rainfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>95&quot;</td>
</tr>
<tr>
<td>2002</td>
<td>118&quot;</td>
</tr>
<tr>
<td>2003</td>
<td>85&quot;</td>
</tr>
<tr>
<td>2004</td>
<td>154&quot;</td>
</tr>
<tr>
<td>2005</td>
<td>102&quot;</td>
</tr>
<tr>
<td>2006</td>
<td>96&quot;</td>
</tr>
</tbody>
</table>
Guyana rainfall data

- We have data that can be modeled as coming from a **Normal distribution** with **mean** $\mu$ and **standard deviation** $\sigma$ (both unknown).

- In **Guyana, South America**, we record the **annual rainfall in the last 6 years**.

<table>
<thead>
<tr>
<th>Year</th>
<th>rainfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>95&quot;</td>
</tr>
<tr>
<td>2002</td>
<td>118&quot;</td>
</tr>
<tr>
<td>2003</td>
<td>85&quot;</td>
</tr>
<tr>
<td>2004</td>
<td>154&quot;</td>
</tr>
<tr>
<td>2005</td>
<td>102&quot;</td>
</tr>
<tr>
<td>2006</td>
<td>96&quot;</td>
</tr>
</tbody>
</table>
Normal likelihood

- Normal distribution is **continuous**. So, probability of observing a **specific value** does not make sense!
- Instead, we have some round-off error.
- In fact, 95" is anything between 94.5" and 95.5".
- So, the actual probability is the integral

\[
\int_{94.5}^{95.5} f(x) \, dx \approx f(95)
\]

So, we take \( f(95) \) as the **approximate** probability.
Normal likelihood

- Normal distribution is continuous. So, probability of observing a specific value does not make sense!
- Instead, we have some round-off error.
- In fact, 95" is anything between 94.5" and 95.5".
- So, the actual probability is the integral

\[ \int_{94.5}^{95.5} f(x) \, dx \approx f(95) \]

So, we take \( f(95) \) as the approximate probability.
Normal likelihood

- Normal distribution is **continuous**. So, **probability** of observing a **specific value** does not make sense!
- Instead, we have some **round-off error**.
- In fact, 95" is anything between 94.5" and 95.5".
- So, the actual probability is the integral

\[ \int_{94.5}^{95.5} f(x) \, dx \approx f(95) \]

So, we take \( f(95) \) as the **approximate probability**.
Normal likelihood

- Also, the rainfall amounts in different years are considered as independent.
- Caution: this is slightly dubious. Ideally we should treat them as time-series but that is beyond our discussion here.
- So, the likelihood of the data is $f(95).f(118)\ldots f(96)$.
- For a specific $\mu$ and $\sigma$, the likelihood is

$$L(\mu, \sigma) = f(95, \mu, \sigma).f(118, \mu, \sigma)\ldots f(96, \mu, \sigma)$$
Normal likelihood

- Also, the rainfall amounts in different years are considered as independent.
- **Caution**: this is slightly dubious. Ideally we should treat them as time-series but that is beyond our discussion here.
- So, the likelihood of the data is \( f(95)f(118) \ldots f(96) \).
- For a specific \( \mu \) and \( \sigma \), the likelihood is

\[
L(\mu, \sigma) = f(95, \mu, \sigma)f(118, \mu, \sigma) \ldots f(96, \mu, \sigma)
\]
Normal likelihood

Also, the rainfall amounts in different years are considered as independent.

**Caution**: this is slightly dubious. Ideally we should treat them as time-series but that is beyond our discussion here.

So, the likelihood of the data is \( f(95) \cdot f(118) \ldots f(96) \).

For a specific \( \mu \) and \( \sigma \), the likelihood is

\[
L(\mu, \sigma) = f(95, \mu, \sigma) \cdot f(118, \mu, \sigma) \ldots f(96, \mu, \sigma)
\]
Plot using MATLAB

- We can plot that function using **MATLAB**.
- For any $\mu$, $\sigma$ and vector $X$, the function $\text{normpdf}(X,\mu,\sigma)$ returns a vector of f-values.
- Likelihood is a product of those values.

```matlab
clear all
X = [95,118, 85,154,102,96]';
mu = 100;
sigma = 10;
L(mu,sigma) = prod(normpdf(X,mu,sigma))
```
We can plot that function using **MATLAB**.

For any $\mu, \sigma$ and vector $X$, the function $\text{normpdf}(X, \mu, \sigma)$ returns a vector of f-values.

Likelihood is a product of those values.

```matlab
clear all
X = [95,118, 85,154,102,96]';
mu = 100;
sigma = 10;
L(mu,sigma) = prod(normpdf(X,mu,sigma))
```
Plot using MATLAB

- We can plot that function using MATLAB.
- For any $\mu$, $\sigma$ and vector $X$, the function \( \text{normpdf}(X,\mu,\sigma) \) returns a vector of f-values.
- Likelihood is a product of those values.

```matlab
clear all
X = [95, 118, 85, 154, 102, 96]';
mu = 100;
sigma = 10;
L(mu,sigma) = prod(normpdf(X,mu,sigma))
```
Possible values of $\mu$ and $\sigma$

- The data are between 85 and 154.
- So, $\mu$, as a measure of central tendency, should be between these values.
- Range of the data is 154-85=69.
- $\sigma$ is likely to be between 0 and 70.
- We plot $L(\mu, \sigma)$ for those values.
Possible values of $\mu$ and $\sigma$

- The data are between 85 and 154.
- So, $\mu$, as a measure of central tendency, should be between this values.
- **Range** of the data is 154-85=69.
- $\sigma$ is likely to be between 0 and 70.
- We plot $L(\mu, \sigma)$ for those values.
Possible values of $\mu$ and $\sigma$

- The data are between 85 and 154.
- So, $\mu$, as a measure of central tendency, should be between this values.
- Range of the data is 154-85=69.
- $\sigma$ is likely to be between 0 and 70.
- We plot $L(\mu, \sigma)$ for those values.
Plot of the function $L(\mu, \sigma)$

MATLAB code:

```matlab
mu = [85:.3:154];
sigma = [0:.5:70];
L = zeros(length(mu),length(sigma));
for i = 1:length(mu)
    for j = 1:length(sigma)
        L(i,j) = prod(normpdf(X,mu(i),sigma(j)));
    end
end

surf(mu,sigma, L')
xlabel('\sigma')
ylabel('\mu')
```
Plot of $L(\mu, \sigma)$
Computation of the MLE-s

\[ L(\mu, \sigma) = f(95, \mu, \sigma) \cdot f(118, \mu, \sigma) \cdots f(96, \mu, \sigma) \]

Recall that

\[ f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi \sigma}} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) \]

Therefore

\[ L(\mu, \sigma) = \frac{1}{8\pi^3 \sigma^6} \exp \left( -\frac{1}{2\sigma^2} \left\{ (95 - \mu)^2 + \ldots + (96 - \mu)^2 \right\} \right) \]

\[ = \frac{1}{8\pi^3 \sigma^6} \exp \left( -\frac{1}{2\sigma^2} \left\{ 6\mu^2 - 1300\mu + 73510 \right\} \right) \]
Computation of the MLE-s

\[ L(\mu, \sigma) = f(95, \mu, \sigma) \cdot f(118, \mu, \sigma) \cdots f(96, \mu, \sigma) \]

Recall that

\[ f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \]

Therefore

\[
L(\mu, \sigma) = \frac{1}{8\pi^3\sigma^6} \exp\left(-\frac{1}{2\sigma^2}\left\{(95 - \mu)^2 + \cdots + (96 - \mu)^2\right\}\right) \\
= \frac{1}{8\pi^3\sigma^6} \exp\left(-\frac{1}{2\sigma^2}\{6\mu^2 - 1300\mu + 73510\}\right)
\]
Computation of the MLE-s

\[ L(\mu, \sigma) = f(95, \mu, \sigma) \cdot f(118, \mu, \sigma) \ldots f(96, \mu, \sigma) \]

Recall that

\[ f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \]

Therefore

\[ L(\mu, \sigma) = \frac{1}{8\pi^3\sigma^6} \exp\left(-\frac{1}{2\sigma^2}\left\{(95 - \mu)^2 + \ldots + (96 - \mu)^2\right\}\right) \]
\[ = \frac{1}{8\pi^3\sigma^6} \exp\left(-\frac{1}{2\sigma^2}\left\{6\mu^2 - 1300\mu + 73510\right\}\right) \]
Computation of the MLE-s

- Taking logarithm

\[ l(\mu, \sigma) = C - 6 \log \sigma - \frac{1}{\sigma^2} \{3\mu^2 - 650\mu + 36755\} \]

- Taking partial derivatives,

\[ \frac{\partial l}{\partial \mu} = - \frac{1}{\sigma^2} \{6\mu - 650\} \]

and

\[ \frac{\partial l}{\partial \sigma} = - \frac{6}{\sigma} + \frac{2}{\sigma^3} \{3\mu^2 - 650\mu + 36755\} \]

- Setting \( \frac{\partial l}{\partial \mu} = \frac{\partial l}{\partial \sigma} = 0 \), we get

\[ \mu = 108.33, \quad \sigma = 22.7061 \]
Computation of the MLE-s

- Taking logarithm

\[ l(\mu, \sigma) = C - 6 \log \sigma - \frac{1}{\sigma^2} \{3\mu^2 - 650\mu + 36755\} \]

- Taking partial derivatives,

\[
\frac{\partial l}{\partial \mu} = -\frac{1}{\sigma^2} \{6\mu - 650\}
\]

and

\[
\frac{\partial l}{\partial \sigma} = -\frac{6}{\sigma} + \frac{2}{\sigma^3} \{3\mu^2 - 650\mu + 36755\}
\]

- Setting \( \frac{\partial l}{\partial \mu} = \frac{\partial l}{\partial \sigma} = 0 \), we get

\[ \mu = 108.33, \quad \sigma = 22.7061 \]
Computation of the MLE-s

- Taking logarithm

\[ l(\mu, \sigma) = C - 6 \log \sigma - \frac{1}{\sigma^2} \{3\mu^2 - 650\mu + 36755\} \]

- Taking partial derivatives,

\[ \frac{\partial l}{\partial \mu} = - \frac{1}{\sigma^2} \{6\mu - 650\} \]

and

\[ \frac{\partial l}{\partial \sigma} = - \frac{6}{\sigma} + \frac{2}{\sigma^3} \{3\mu^2 - 650\mu + 36755\} \]

- Setting \( \frac{\partial l}{\partial \mu} = \frac{\partial l}{\partial \sigma} = 0 \), we get

\[ \mu = 108.33, \quad \sigma = 22.7061 \]
Exercise 2

- Enter the data on **heights** (as collected by Dhruv).
- Assume that the data comes from a **normal distribution**.
- Plot the **likelihood** as a function of $\mu$ and $\sigma$. (Use the range of the data for the range of $\mu$ and $\sigma$.)
- Find the **mean and the standard deviation** of the data. (in MATLAB use the functions `mean(x)` and `std(x,1)`)
- Check (visually), that the **mean and SD** corresponds for the **MLE** for $\mu$ and $\sigma$ respectively.
Exercise 2

- Enter the data on heights (as collected by Dhruv).
- Assume that the data comes from a normal distribution.
- Plot the likelihood as a function of $\mu$ and $\sigma$. (Use the range of the data for the range of $\mu$ and $\sigma$.
- Find the mean and the standard deviation of the data. (in MATLAB use the functions `mean(x)` and `std(x,1)`)
- Check (visually), that the mean and SD corresponds for the MLE for $\mu$ and $\sigma$ respectively.
Exercise 2

- Enter the data on heights (as collected by Dhruv).
- Assume that the data comes from a normal distribution.
- Plot the likelihood as a function of $\mu$ and $\sigma$. (Use the range of the data for the range of $\mu$ and $\sigma$.
- Find the mean and the standard deviation of the data. (in MATLAB use the functions $\text{mean}(x)$ and $\text{std}(x,1)$)
- Check (visually), that the mean and SD corresponds for the MLE for $\mu$ and $\sigma$ respectively.
Exercise 2

- Enter the data on heights (as collected by Dhruv).
- Assume that the data comes from a normal distribution.
- Plot the likelihood as a function of $\mu$ and $\sigma$. (Use the range of the data for the range of $\mu$ and $\sigma$.
- Find the mean and the standard deviation of the data. (in MATLAB use the functions `mean(x)` and `std(x,1)`)
- Check (visually), that the mean and SD corresponds for the MLE for $\mu$ and $\sigma$ respectively.
Conclusion

- As shown in the examples, the **MLE of (multiple) parameters** can be done simultaneously.
- In both examples, there is an **utility function** that we need to maximize.
- Similarly, there may be a **cost function** attached to the parameters that we can **minimize** to get estimators.
- The least squares estimation or the least absolute deviation methods are from that class of estimation.

Enjoy your stay here in Raleigh !!!
Conclusion

■ As shown in the examples, the MLE of (multiple) parameters can be done simultaneously.

■ In both examples, there is an utility function that we need to maximize.

■ Similarly, there may be a cost function attached to the parameters that we can minimize to get estimators.

■ The least squares estimation or the least absolute deviation methods are from that class of estimation.

Enjoy your stay here in Raleigh !!!
Conclusion

- As shown in the examples, the MLE of (multiple) parameters can be done simultaneously.
- In both examples, there is an utility function that we need to maximize.
- Similarly, there may be a cost function attached to the parameters that we can minimize to get estimators.
- The least squares estimation or the least absolute deviation methods are from that class of estimation.

Enjoy your stay here in Raleigh !!
Conclusion

- As shown in the examples, the **MLE of (multiple) parameters** can be done simultaneously.
- In both examples, there is an **utility function** that we need to maximize.
- Similarly, there may be a **cost function** attached to the parameters that we can **minimize** to get estimators.
- The **least squares estimation** or the **least absolute deviation methods** are from that class of estimation.

Enjoy your stay here in Raleigh !!