To Do List

Read Chapter 2

Do Problems 1- 11 in Chapter 2

Last Class: Parameter Estimation

1) Definition of a (Point) Estimate of an Unknown Parameter

2) Method of Maximum Likelihood

i) Definition of the Likelihood Function

ii) Definition of the Maximum Likelihood Estimate

iii) Definition of the Relative Likelihood Function

iv) Definition of the Log Likelihood Function

Today's Lecture (Sec. 2.2 - 2.5):

1) Likelihood Function for a Random Sample

2) Likelihood Functions for Continuous Random Variables (Exponential & Gaussian)

3) Likelihood Function for Multinomial Model

Likelihood Function for a Random Sample (General Case)

Suppose the data are of the form $y = (y_1, y_2, \ldots, y_n)$ where $(y_1, y_2, \ldots, y_n)$ is assumed to be a realization of the random vector $Y = (Y_1, Y_2, \ldots, Y_n)$.

Suppose also that the Y_i's are assumed to be independent and identically distributed (i.i.d) random variables with probability function

$$P(Y=y;\ \theta) = f(y;\ \theta),\ \theta \in \Omega.$$

$Y_1, Y_2, \ldots, Y_n$ is called a random sample.

Likelihood Function for a Random Sample

The likelihood function for $\boldsymbol{\theta}$ based on the observed random sample $\boldsymbol{y_1, y_2, \ldots, y_n}$ is

$$L(\theta) = P(\text{observing the data } y_1, y_2, \ldots, y_n; \theta)$$

$$= P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n; \theta)$$

$$= P(Y_1 = y_1; \theta) P(Y_2 = y_2; \theta) \cdots P(Y_n = y_n; \theta)$$

$$\text{since the } Y_i\text{'s are independent r.v.'s}$$

$$= \prod_{\mathrm{i=1}}^{\mathrm{n}} P(Y_i = y_i; \theta) \qquad \theta \in \Omega.$$

Likelihood Functions for Continuous Random Variables

For discrete random variables the likelihood function is equal to the probability of observing the data y, that is,

$$L(\theta) = L(\theta; \boldsymbol{y}) = P(\boldsymbol{Y} = \boldsymbol{y}; \theta) \text{ for } \theta \in \Omega$$

$$= \text{the probability that we observe the data } \boldsymbol{y} \text{ as a function of } \theta$$

For continuous distributions, this is unsuitable as a definition for *L*(*θ*) since P(*Y* = *y*; *θ*) = 0.

Likelihood Functions for Continuous Random Variables Cont'd

Suppose $Y = (Y_1, Y_2, \ldots, Y_n)$ is a random sample from a continuous distribution with probability density function $f(y;\theta)$ for $\theta \in \Omega$.

Suppose also that $y = (y_1, y_2, \ldots, y_n)$ represents a realization of $Y = (Y_1, Y_2, \ldots, Y_n)$.

We define the likelihood function for θ based on the observed data $y = (y_1, y_2, \ldots, y_n)$ as

$$L(\theta) = L(\theta; \boldsymbol{y}) = \prod_{i=1}^{n} f(y_i; \theta) \qquad \theta \in \Omega.$$

Likelihood Function for Exponential Distribution (Course Notes p. 57)

Suppose Y = lifetime of a randomly selected light bulb in a large population and that $Y \sim$ Exponential(θ).

A random sample of light bulbs is tested and the lifetimes $y_1, y_2, \ldots, y_n$ are observed.

Find the maximum likelihood estimate of θ based on the observed data $y_1, y_2, \ldots, y_n$.

Exponential Likelihood Function

For Exponential data $y_1, y_2, \ldots, y_n$ the likelihood function for θ (ignoring constants with respect to θ) is

$$L(\theta) = \theta^{-n} e^{-n\bar{y}/\theta} \quad for \;\; \theta > 0$$

and the maximum likelihood estimate of θ is

$$\hat{\theta} = \bar{y}$$

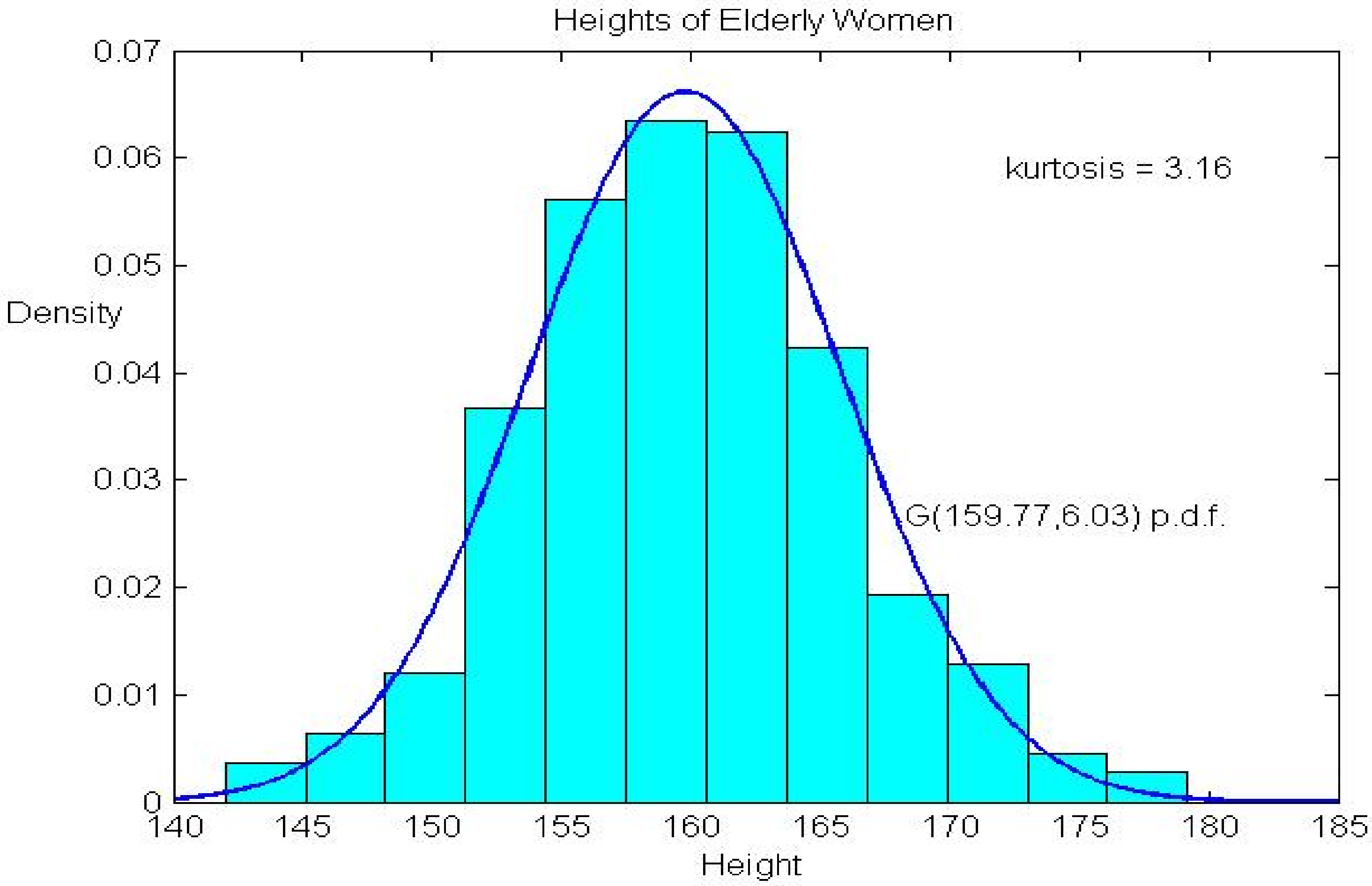
Likelihood Function for Gaussian Data

Recall previous example on heights of elderly women in a study of osteoporosis.

Example: Heights in centimeters of a sample of 351 elderly women randomly selected from a community in a study of osteoporosis.

156	163	169	161	154	156	163	164	156	166	177	158
150	164	159	157	166	163	153	161	170	159	170	157
156	156	153	178	161	164	158	158	162	160	150	162
155	161	158	163	158	162	163	152	173	159	154	155
164	163	164	157	152	154	173	154	162	163	163	165
160	162	155	160	151	163	160	165	166	178	153	160
156	151	165	169	157	152	164	166	160	165	163	158
153	162	163	162	164	155	155	161	162	156	169	159
159	159	158	160	165	152	157	149	169	154	146	156
157	163	166	165	155	151	157	156	160	170	158	165
167	162	153	156	163	157	147	163	161	161	153	155
166	159	157	152	159	166	160	157	153	159	156	152
151	171	162	158	152	157	162	168	155	155	155	161
157	158	153	155	161	160	160	170	163	153	159	169
155	161	156	153	156	158	164	160	157	158	157	156
160	161	167	162	158	163	147	153	155	159	156	161
158	164	163	155	155	158	165	176	158	155	150	154
164	145	153	169	160	159	159	163	148	171	158	158
157	158	168	161	165	167	158	158	161	160	163	163
169	163	164	150	154	165	158	161	156			
154	158	162	164	158	165	158	156	162			
157	167	142	166	163	163	151	163	153			
169	154	155	167	164	170	174	155	157			
155	168	152	165	158	162	173	154	167			
158	167	164	170	164	166	170	160	148			
150	165	165	147	162	165	158	145	150			
163	166	162	163	160	162	153	168	163			
158	155	168	160	153	163	161	145	161			
161	155	158	161	163	157	156	152	156			
160	152	153									

Heights of Elderly Women Data



$$\bar{y} = 159.77, \quad s^2 = 36.36, \quad s = 6.03$$

Model for Heights of Elderly Women

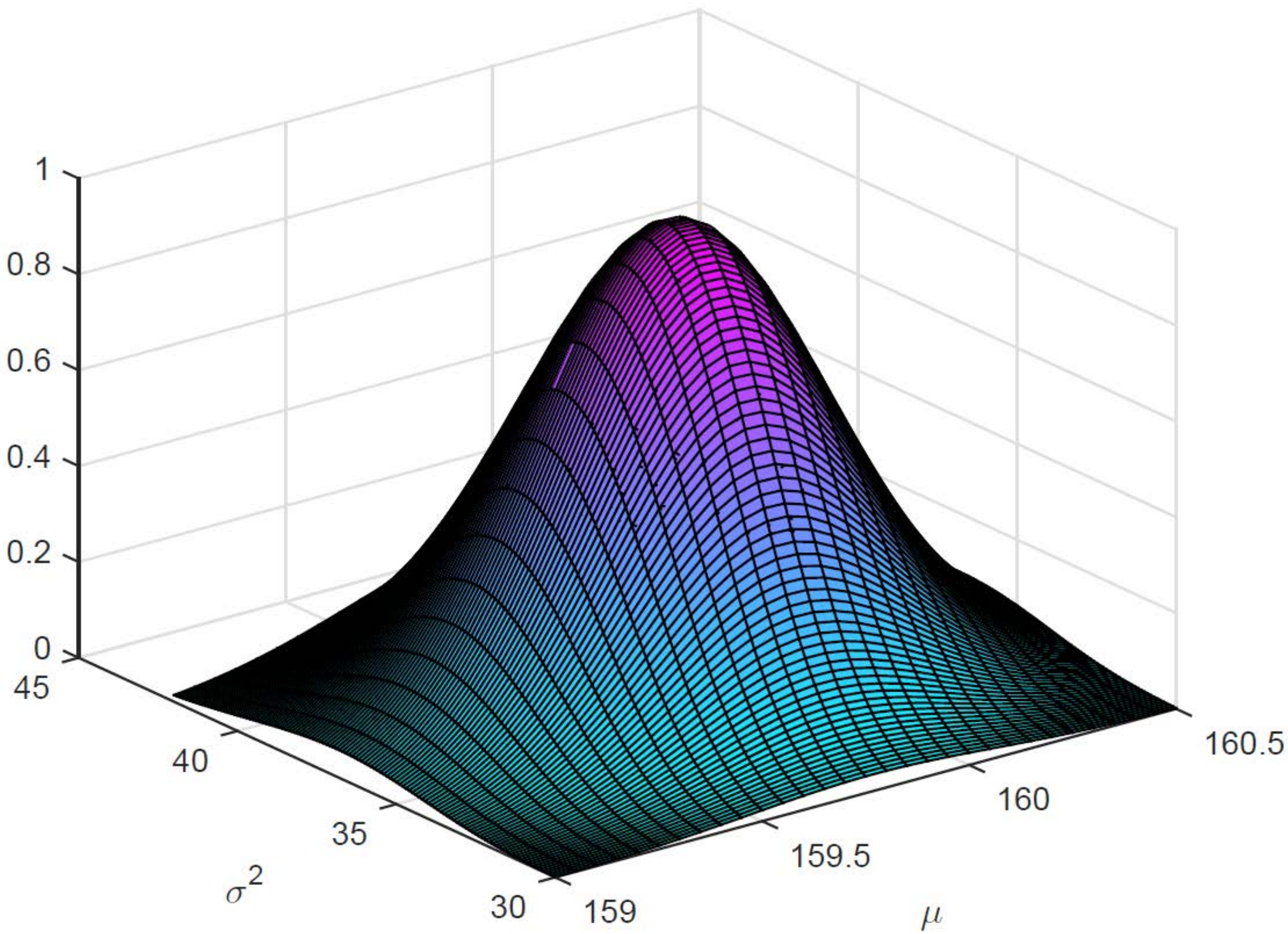
Given the agreement between the relative frequency histogram and the Gaussian probability density function it seems reasonable to assume the following model for the heights of elderly women:

Let Y = height of a randomly selected elderly woman in a large population and that $Y \sim G(\mu,\sigma)$.

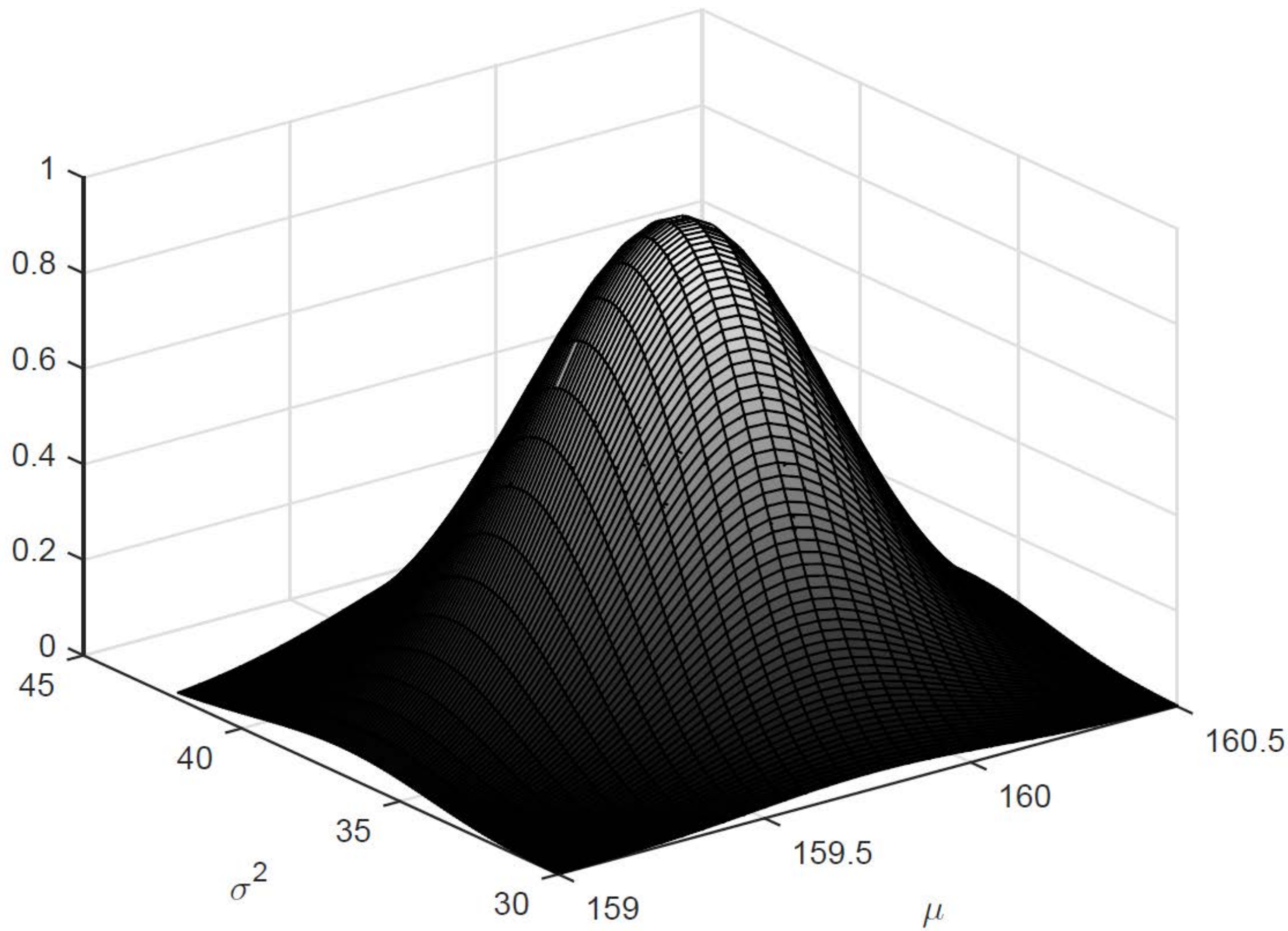
A random sample of elderly women is select and their heights $y_1, y_2, \ldots, y_n$ are observed.

What is the maximum likelihood estimate of $\boldsymbol{\theta} = (\mu, \sigma^2)$?

Gaussian Likelihood Function



50 Shades of Grey



Maximum Likelihood Estimates for Gaussian Data

$$\hat{\mu} = \overline{y} \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2$$

See further derivation details in Example 2.3.2 on page 58 of the Course Notes.

Data from Previous Class

Indicate your agreement with the following statement:
"It is important for all students in the Faculty of Mathematics to take a course in Introductory Statistics."

A: Strongly Agree

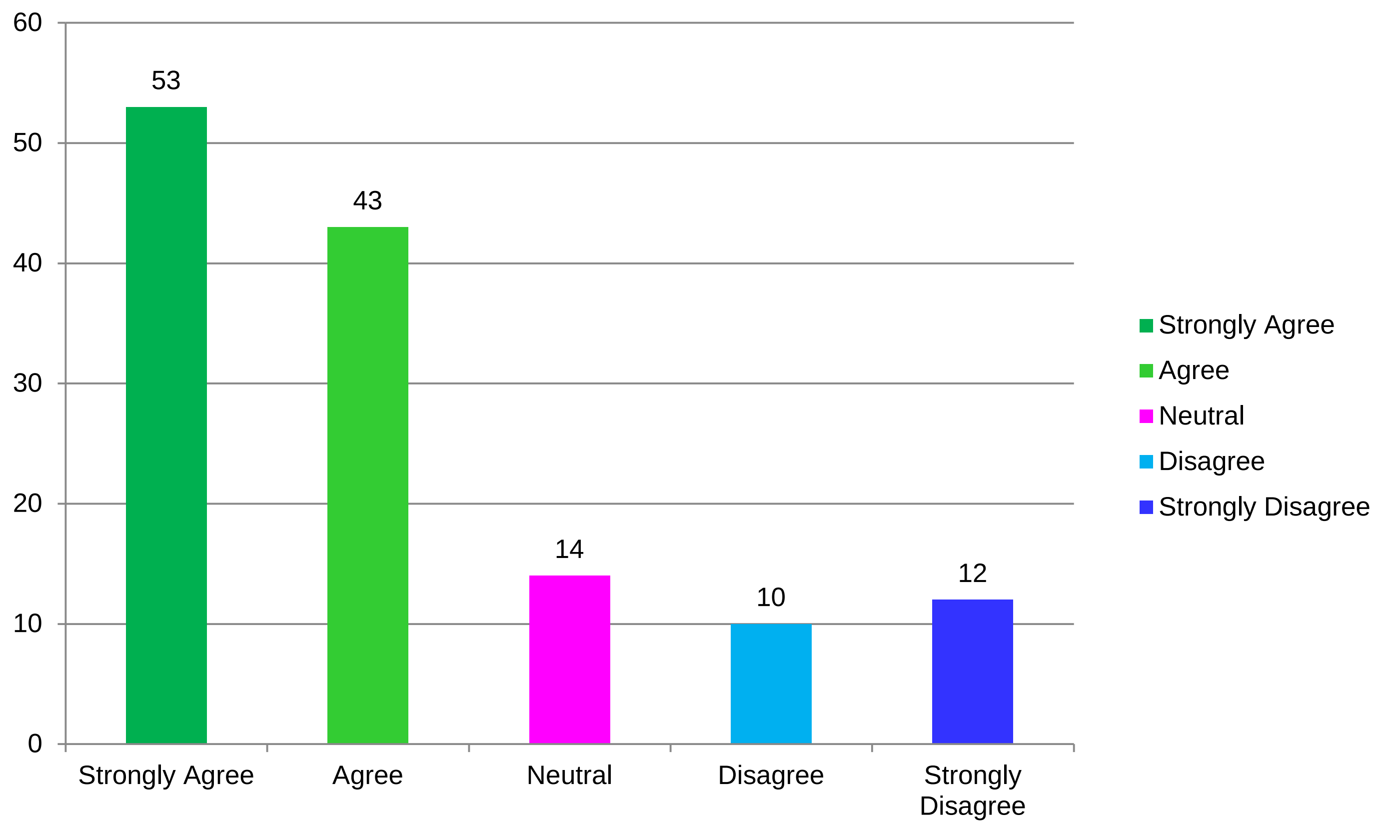
B: Agree

C: Neutral

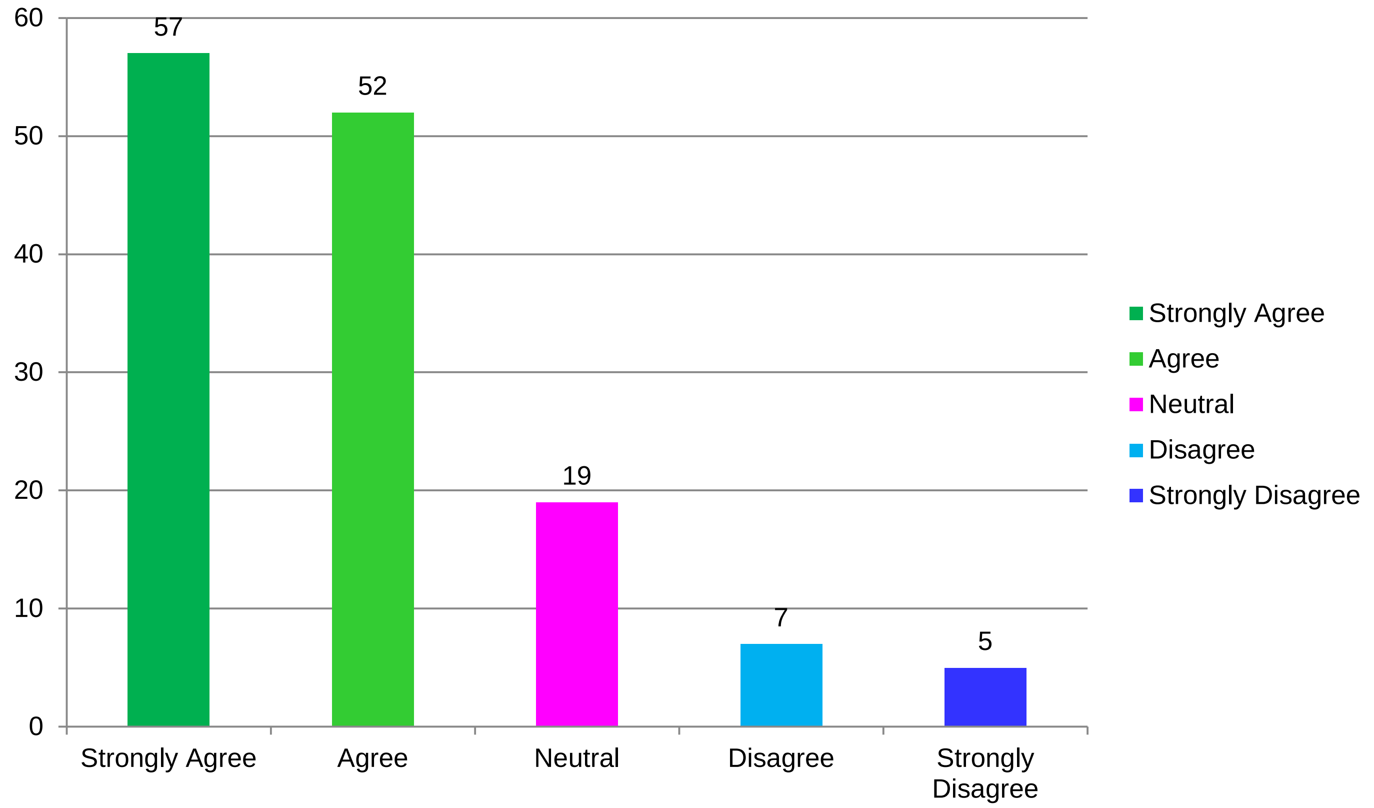
D: Disagree

E: Strongly Disagree

Bar Graph for Fall 2016 Data



Bar Graph for Winter 2016 Data



Likelihood Function for Multinomial Distribution (Sec. 2.4)

Let the random variables Y_1, Y_2, Y_3, Y_4 , Y_5 represent the number who "Strongly Agree", "Agree", "Neutral", "Disagree", and "Strongly Disagree" that we might observe in a random sample of size *n* people.

Then

(Y_1, Y_2, Y_3, Y_4 , Y_5)
$\sim$ Multinomial(n ; θ_1, θ_2, θ_3, θ_4, θ_5).

Maximum Likelihood Estimates Fall 2016

The maximum likelihood estimates of θ_1, θ_2, θ_3, θ_4, θ_5 from the observed Fall 2016 data are:

$$\hat{\theta}_1 = \frac{53}{132} = 0.40, \qquad \hat{\theta}_2 = \frac{43}{132} = 0.33,$$

$$\hat{\theta}_3 = \frac{14}{132} = 0.11, \qquad \hat{\theta}_4 = \frac{10}{132} = 0.08, \qquad \hat{\theta}_5 = \frac{12}{132} = 0.09,$$

$$\text{Note:} \quad \hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3 + \hat{\theta}_4 + \hat{\theta}_5 = 1.$$

Maximum Likelihood Estimates Fall 2016

The maximum likelihood estimates of $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$, $\boldsymbol{\theta_3}$, $\boldsymbol{\theta_4}$, $\boldsymbol{\theta_5}$ from the Winter 2016 data are:

$$\hat{\theta}_1 = \frac{57}{140} = 0.41, \quad \hat{\theta}_2 = \frac{52}{140} = 0.37,$$

$$\hat{\theta}_3 = \frac{19}{140} = 0.14, \quad \hat{\theta}_4 = \frac{7}{140} = 0.05, \quad \hat{\theta}_5 = \frac{5}{140} = 0.04,$$

$$\text{Note:} \quad \hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3 + \hat{\theta}_4 + \hat{\theta}_5 = 1.$$

Multinomial Model

Please look at the other examples in Section 2.4 of the Course Notes.

We will look more closely at examples based on the Multinomial model in Chapter 7.