Midterm Study Guide

Analysis of Variance

Quinton Quagliano, M.S., C.S.P

Department of Educational Psychology

Table of Contents

1 Introduction

This is the instructor-provided study guide for the midterm exam for EDPS-641. Please use this document as **one** of the aids in preparing for this exam, but I strongly encourage you to develop additional study aids and tools on your own. This document is **not meant to be comprehensive** to all of the content covered in modules 1 through 7 - it is a merely a summary of what was covered, and meant as a review tool to help guide your practice.

Midterm / Exam 1 Structure covers the basic policies and procedures of the test, with information lifted directly from the syllabus. Tips for Preparing highlights what I feel are good practices to employ as you get ready to take the test. Tips for Studying contains ideas on what additional steps you can take to help get ready for the exam. Vocabulary contains many (but not necessarily all) of the important terms and phrases used in the first 7 modules, and Formulas and Notation will give a brief overview of many of the formulas you have been introduced to across the modules.

Please do feel free to reach out with any questions you may have, but also make ready use of these many opportunities and resources to do well. I believe in each of you to use your time and diligence to turn in an excellent performance on this test.

2 Midterm / Exam 1 Structure

Please review the following information and procedures regarding the midterm and final exams in this course. The following text is also in the syllabus:

There will be 2 conceptual exams in this course, effectively a midterm and a final. These conceptual exams are intended to be **cumulative** and will cover content from all covered units and weekly quizzes. Much like the weekly quizzes, exams will not be focused on calculation of statistics, but rather, on a conceptual understanding (i.e., you shouldn't have to do any math).

The format is as follows:

- Each exam is 40 multiple-choice questions, 1 points for each question
- Exams will be taken on the Canvas LMS
- Exams will contain content from the entire unit, between all lectures AND readings and any other activities
- Exams are open-note, you may use the "skeleton notes" that I provide, or your own written notes. Thus, the exams reward good structure in thoughtfulness in your notes and preparation
- You may not collaborate with others during the exam, or discuss questions with other students after the exam. You cannot use AI tools, the internet, or any electronic devices to help you. You may not use the book or slides, only printed/handwritten notes.

- Exams will be graded promptly and reviewed the following week; the correct answers to the questions will also be provided after all students have had an opportunity to take the exam

3 Tips for Preparing

The below pieces of advice are my opinionated thoughts about how you may successfully prepare for this exam. Each student may find different strategies and techniques useful, please customize the following advice to your own needs.

- Because the exam is “open note”, it is **extremely** important that you gather all relevant notes you may need prior to beginning the exam. While it is not timed, you should complete it in entirely one sitting. I recommend having access to:
  - Any and all lecture notes taken, whether that be the guided notes filled out or your own personal notes
  - Printed copies of your lecture check-ins, annotated answer keys from me
  - Printed off copies of any practical assignment(s) if you find the applied examples more elucidating
  - Any additional handwritten or student-typed preparation, such as answers to this study guide
- While the structure of this exam is loose, you **should** still study as if this will be a timed, secure test - you should not need to look up the answer to each question, as that would be too time-consuming
- You should aim to be finished with the test in roughly an hour, but you are allowed as much time
- You should take the exam on a laptop or desktop computer, not a tablet or phone. You should have your computed plugged in during the exam to reduce risk of battery running out, and should warn others you live with to not disturb you until you are done
- Find a quiet, secluded spot with good internet connection and an outlet for your computer.
- Prepare any necessary snack(s) and drink(s) for when you complete the exam, so you do not have to step away from your work
- Place any other devices and distractions (e.g., phone, earbuds, etc.) in another room, unless necessary for medical or academic accommodations

4 Tips for Studying

The below pieces of advice are my opinionated thoughts about how you may successfully study for this exam. Each student may find different strategies and techniques useful, please customize the following advice to your own needs.

- As a general rule: **create, don't consume!** The best study materials are the ones that you make and develop. While I give you this document as a starting point, you should make the notes and study materials that are best suited to *you*.
- **Give yourself enough time!** Good studying starts early, and remains consistent. There is simply too much material to making cramming an effective strategy.
- **DO NOT use this study guide as your sole resource when studying.** I provided this guide as a convenience, but I make no certain guarantee that it contains any and all information relevant to the exam. Any of the content covered in lectures or reading is fair game to be tested upon - and you should use resources from those things to prepare. Completion of this study guide does not necessarily guarantee success on the actual exam.
- If you do not know where to start:
  - First, focus on the Vocabulary and being able to organically come up with definitions and explanations for those terms. This is where many students may opt to use flashcards or spaced repetition software like Quizlet or Anki.
  - Then, focus on the Formulas and Notation. You do not need to fully memorize each one, but you may benefit from fast recognition, i.e., "Oh that looks like a probability density function". Try to work through describing each one practically - why are certain parts there and what is happening in the formula? When might certain formulas show up, and what other ones are they related to?
- Try "teaching" the materials to an (un)willing volunteer or inanimate object
  - Trying to teach through and explain a concept will immediately make it apparent whether you understand it, or do not. This is how I double-check myself when writing lectures: if I can't explain something fluidly, then I haven't truly mastered explaining it well enough yet.
  - You could even try using my slides and see if you can talk "between the lines" and go beyond what is just written down - lecturing on a certain piece of content can actually really help you reinforce your understanding.
- Review past quizzes and lecture check-in answers; ensure you have corrected your understanding where you may have gotten a question wrong the first time through
- Consider "re-annotating" your previous notes:
  - Add question marks where you are confused on something
  - Add highlights to parts that stick out as especially critical to understanding
  - Re-work discussion and multiple choice questions strewn throughout and make sure you have good answers for each of them

- Review learning objectives posted at the start of each lecture, and assess whether you feel comfortable in meeting those objectives
- Complete practice and review questions at the end of each chapter of the textbook - while many focus on calculations, working those processes might very well see how the Formulas and Notation below are used in practice.
- Consider coming to office hours for the graduate assistant or the instructor to clarify difficult topics

5 Vocabulary

The terms in this section have been lifted from the textbook bolded terms and may not have been used directly in lecture, though I have tried to remove those not explicitly covered by my recordings. If you cannot find a clear definition or example, consider looking at the end of each chapter for the author-provided definition, or review the video from the respective module to see my explanation again.

Some terms may be duplicated due to having appeared as bolded terms in multiple chapters, consider their relevance to the context of each module.

5.1 Module 1

Average

Categorical Variable

Cluster Sampling

Continuous Random Variable

Convenience Sampling

Cumulative Relative Frequency

Data

Frequency

Nonsampling Error

Numerical Variable

Parameter

Population

Probability

Proportion

Qualitative Data

Quantitative Data

Random Assignment

Random Sampling

Relative Frequency

Representative Sample

Response Variable

Sample

Sampling Bias

Sampling Error

Sampling with Replacement

Sampling without Replacement

Simple Random Sampling
Statistic
Stratified Sampling
Systematic Sampling
Variable

5.2 Module 2

Box plot
First Quartile
Frequency
Frequency Polygon
Frequency Table
Histogram
Interquartile Range
Interval
Mean
Median
Midpoint
Mode
Outlier
Percentile
Quartiles
Relative Frequency
Skewed
Standard Deviation
Variance

5.3 Module 3

AND Event
Complement Event
Conditional Probability
Conditional Probability of A GIVEN B
Conditional Probability of One Event Given Another Event
Contingency table
Dependent Events
Equally Likely
Event
Experiment
Independent Events
Mutually Exclusive
Or Event
Outcome
Probability
Sample Space
Tree Diagram
Venn Diagram

5.4 Module 4

Bernoulli Trials

Binomial Experiment

Binomial Probability Distribution

Expected Value

Mean

Mean of a Probability Distribution

Probability Distribution Function (PDF)

Random Variable (RV)

Standard Deviation of a Probability Distribution

The Law of Large Numbers

5.5 Module 5

Conditional Probability

Decay parameter

Exponential Distribution

Poisson distribution

Uniform Distribution

5.6 Module 6

Normal Distribution

Standard Normal Distribution

Z-score

5.7 Module 7

Average

Central Limit Theorem

Exponential Distribution

Mean

Normal Distribution

Normal Distribution

Sampling Distribution

Standard Error of the Mean

Uniform Distribution

6 Formulas and Notation

6.1 General Information

Formulas and notation are often the scariest part of statistics for many folks, and though we won't have to calculate them during the exam, I still expect you to recognize and understand them.

Some formulas and notation are given both for sample and population (be careful of small differences between the two), and sometimes there may be two alternative formulas listed for the same thing, in the case there is an equivalent form. In writing the test, I will stick to the notations shown here in the study guide and in the lecture notes.

Some notation has different meanings depending on the formula it is part of, e.g., a population mean and expected long-term mean have the same notation of μ, but have different meanings and context. Please be mindful of navigating the formulas.

Like the rest of the content in this study guide, this section is not necessarily exhaustive, and you should review your notes and the textbook for additional context to each of these.

6.2 Module 1

No formulas introduced in this module

6.3 Module 2

6.3.1 Inter-quartile Range (IQR)

Used for describing the middle 50% of the data, i.e., that which lies between the first and the third quartiles

$$IQR = Q_3 - Q_1$$

Where:
- Q_1 is the first quartile
- Q_2 is the second quartile
- Q_3 is the third quartile

6.3.2 Finding Index of Value at Kth Percentile

Used for finding which value in a given dataset is closest to a specified percentile we are interested in

$$i = \frac{k}{100} * (n + 1)$$

Where:
- i is the index or rank of the value when ordered smallest to largest
- k is the k^{th} percentile

- n is the total number of data points

6.3.3 Finding Percentile of Value

Used to find the rough percentile of a certain point in a dataset,

$$\% = \frac{x + 0.5y}{n} * 100$$

Where:
- x is the number of data points up to and NOT including the point of interest
- y the number of occurrences of the values of interest
- n is the total number of data points

6.3.4 Mean / Average

Used to find the arithmetic mean (also known as average) of a dataset

Sample:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$$

Population:

$$\mu = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$$

Where:
- x_1 is the first value in the data, x_2 is the second value, and so on, until x_n is the final number in the data
- n is the total number of data points

6.3.5 Deviation

Used as a description to say how “far” a data point is away from the mean

Sample:

$$x - \bar{x}$$

$$x_i - \bar{x}$$

Population:

$$x - \mu$$

$$x_i - \mu$$

Where:

- x or x_i is a single value in the data
- $\bar{x}$ is the sample mean
- μ is the population mean

6.3.6 Variation

Usually used as a stepping stone formula to derive standard deviation. Variation is also sometimes described as the averages of the squared deviations, which is intuitive looking at the formula set up and how it combines calculations for the mean and deviations. Worth noting that the squaring of the deviation is necessary to prevent their sum from always being 0, hence why we have to take the square root to get the standard deviation.

Sample:

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

Population:

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

Where:

- x is a single value in the data
- $\bar{x}$ is the sample mean
- μ is the population mean
- n is the total number of data points

6.3.7 Standard Deviation

The standard deviation is the most popular way to describe the general spread of the data from the mean. A larger standard deviation suggests the data is more spread out

Sample:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{s^2}$$

Population:

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{n}}$$

$$\sigma = \sqrt{\sigma^2}$$

Where:

- x is a single value in the data
- $\bar{x}$ is the sample mean
- μ is the population mean
- s^2 is the sample variation
- σ^2 is the population variation
- n is the total number of data points

6.3.8 Z-score

Z-scores are values given to individual data points that represent how many standard deviations they are away from the mean of the data

Sample:

$$z = \frac{x - \bar{x}}{s}$$

Population:

$$z = \frac{x - \mu}{\sigma}$$

Where:
- x is a individual data value of interest
- $\bar{x}$ is the sample mean
- μ is the population mean
- s^2 is the sample variation
- σ^2 is the population variation

6.4 Module 3

6.4.1 Sample Space (List Notation)

The sample space can be represented as a venn diagram or tree as well, but commonly a list helps show every possible outcome from a single probability experiment.

$$S = \{..., ..., ...\}$$

Where:
- Each ... represents one possible outcome from a probability experiment

Example: Coin flip with possible heads (H)/tails (T) outcome: $S = \{H, T\}$

6.4.2 Probability of an Event

We are often interested in some subset of the outcomes in a sample space and we might define which outcomes those are for that specific event. The probability that one of those outcomes occurs is then the probability of that event.

$$P(...)$$

Where:
- ... is some defined event, usually represented by a capital (uppercase) letter

Example: A is an outcome of heads in a fair coin toss, thus, $P(A) = 0.50$

6.4.3 Probability of an OR Event

This is used when we are interested in outcomes that are part of one or more of the selected events.

$$P(... \cup ...)$$

Where:
- Each ... is some defined event, usually represented by a capital (uppercase) letter
- $\cup$ designated a set union of the two events

6.4.4 Probability of an AND Event

This is used when we are interested in outcomes that are part of BOTH described events, not just one or the other

$$P(... \cap ...)$$

Where:
- Each ... is some defined event, usually represented by a capital (uppercase) letter
- $\cap$ designated a set intersection of the two events

6.4.5 Probability of a Conditional Event

This is used when we are interested in the probability that a certain event occurs, given that another event is already true

$$P(..._1|..._2) = \frac{..._1 \bigcap ..._2}{..._2}$$

Where:
- Each ... is some defined event, usually represented by a capital (uppercase) letter
- $|$ is the indicator that the first event prior to the bar is given the second event after the bar
- $\cap$ designated a set intersection of the two events

6.4.6 Probability of a Complement of an Event

This is used when we are interested in the probability of the opposite (or inverse) of a certain event occurring.

$$P(...')$$

Where:
- ... is some defined event, usually represented by a capital (uppercase) letter
- $'$ is an indicator the inverse or complement is taken

6.5 Module 4

6.5.1 Expected Long-term Mean for a Discrete Probability Function

This is the "mean" value that would occur across an infinite number of samples or experiments. May be a decimal, and therefore, not make intuitive sense for a discrete variable.

$$\mu = \sum (x \cdot P(x))$$

Where:

- x is a single value in the data
- $P(x)$ is the probability of x occurring
- $\cdot$ is an indicator of multiplication

6.5.2 Expected Long-term Variance for a Discrete Probability Function

Same as with mean, this would be the variance across infinitely many samples of the probability experiment.

$$\sigma^2 = \sum [(x - \mu)^2 \cdot P(x)]$$

Where:

- x is a single value in the data
- $P(x)$ is the probability of x occurring
- μ is the expected long term mean of the PDF
- $\cdot$ is an indicator of multiplication

6.5.3 Expected Long-term Standard Deviation for a Discrete Probability Function

Same as with mean, this would be the standard deviation across infinitely many samples of the probability experiment.

$$\sigma = \sqrt{\sum [(x - \mu)^2 \cdot P(x)]}$$

$$\sigma = \sqrt{\sigma^2}$$

Where:

- x is a single value in the data

- $P(x)$ is the probability of x occurring
- μ is the expected long term mean of the PDF
- σ^2 is the expected long term variance of the PDF
- $\cdot$ is an indicator of multiplication

6.5.4 Notation for a Binomial Distribution

This is used to describe the core characteristics and construction of a binomial distribution for a discrete variable.

$$... \sim B(n,p)$$

Where:

- ... is some defined discrete random variable, usually represented by a capital (uppercase) letter
- B is an arbitrary indicator of the binomial distribution
- n is number of consecutive, independent trials
- p is the $P(success)$ or probability of success in each experiment

6.5.5 Shortcut Expected Mean for Binomial Distributions

The Binomial distribution has special characteristics that make it possible to avoid the longer expected mean formula in favor of this calculation.

$$\mu = n * p$$

Where:

- n is number of consecutive, independent trials
- p is the $P(success)$ or probability of success in each experiment

6.5.6 Shortcut Expected Variance for Binomial Distributions

See above.

$$\sigma^2 = n * p * q$$

Where:

- n is number of consecutive, independent trials
- p is the $P(success)$ or probability of success in each experiment
- q is the $P(failure)$ or probability of failure in each experiment

6.5.7 Shortcut Expected Standard Deviation for Binomial Distributions

See above.

$$\sigma = \sqrt{n * p * q}$$

Where:
- n is number of consecutive, independent trials
- p is the $P(success)$ or probability of success in each experiment
- q is the $P(failure)$ or probability of failure in each experiment

6.6 Module 5

6.6.1 Probability Density Function for Random Continuous Variables

When working with continuous variables, it is more appropriate to use a line function than a table (as was used with discrete variables). This function draws a line in which the area underneath it is the probability.

$$f(..._1) = ..._2$$

Where:
- $..._1$ is some vector of possible outcomes for a continuous random variable, usually represented as a lowercase letter
- $..._2$ is some equation with $..._1$ that draws a curve or line
- $f()$ is a description to say, "the function of $..._1$"

6.6.2 Notation for a Uniform Distribution

This notation is used to describe the core characteristics of a uniform distributed continuous variable.

$$... \sim U(a, b)$$

Where:
- … is some defined continuous uniform random variable, usually represented by a capital (uppercase) letter
- U is an arbitrary indicator of the uniform distribution
- a is the minimum possible value … can take
- b is the maximum possible value … can take

6.6.3 Probability Density Function for a Uniform Random Continuous Variable

This is a specific extension of the idea in Probability Density Function for Random Continuous Variables, but applied specifically to the uniform case.

$$f(...) = \frac{1}{b - a}$$

Where:

- ... is some vector of possible outcomes for a uniform continuous random variable, usually represented as a lowercase letter
- a is the minimum possible value of ...
- b is the maximum possible value of ...

6.6.4 Shortcut Expected Mean for Uniform Distributions

Like with the easier binomial shortcut formulas for mean, variance, and standard deviation, we can used this for an easy expected mean for uniform variables.

$$\mu = \frac{a + b}{2}$$

Where:

- a is the minimum possible value of the variable of interest
- b is the maximum possible value of the variable of interest

6.6.5 Shortcut Expected Variance for Uniform Distributions

See above.

$$\sigma^2 = \frac{(b - a)^2}{12}$$

Where:

- a is the minimum possible value of the variable of interest
- b is the maximum possible value of the variable of interest

6.6.6 Shortcut Expected Standard Deviation for Uniform Distributions

See above.

$$\sigma = \sqrt{\frac{(b-a)^2}{12}}$$

Where:
- a is the minimum possible value of the variable of interest
- b is the maximum possible value of the variable of interest

6.6.7 Notation for an Exponential Distribution

This notation is used to describe the core characteristics of a exponential distributed continuous variable.

$$... \sim Exp(m)$$

$$... \sim Exp(\lambda)$$

Where:
- ... is some defined exponential continuous random variable, usually represented by a capital (uppercase) letter
- Exp is an arbitrary indicator of the exponential distribution
- m or λ is the decay/rate parameter

6.6.8 Probability Density Function for an Exponential Random Continuous Variable

This is a specific extension of the idea in Probability Density Function for Random Continuous Variables, but applied specifically to the exponential case.

$$f(...) = me^{-m \cdot ...}$$

$$f(...) = \lambda \cdot e^{-\lambda \cdot ...}$$

Where:
- ... is some vector of possible outcomes for an exponential continuous random variable, usually represented as a lowercase letter
- m or λ is the decay/rate parameter

- $\cdot$ is an indicator of multiplication
- e is the scientific constant = 2.71828

6.6.9 Decay/Rate Parameter for Exponential Continuous Variables

The rate parameter is a special descriptor used in the exponential case, which is essential for use in the probability density function and in finding expected mean.

$$m = \frac{1}{\mu}$$

$$\lambda = \frac{1}{\mu}$$

Where:

- μ is the long term expected mean for the exponential variable

6.6.10 Shortcut Expected Mean for Exponential Distributions

Like with the easier uniform shortcut formulas for mean, variance, and standard deviation, we can used this for an easy expected mean for exponential variables.

$$\mu = \frac{1}{m}$$

Where:

- m or λ is the decay/rate parameter

6.6.11 Shortcut Expected Standard Deviation for Exponential Distributions

See above. Since standard deviation is had directly from mean, we don't really *need* a formula for variance in the exponential case.

$$\sigma = \mu$$

Where:

- μ is the long term expected mean for the exponential variable

6.7 Module 6

6.7.1 Notation for a Normal Distribution

This notation is used to describe the core characteristics of a normal distributed continuous variable.

$$... \sim N(\mu, \sigma)$$

Where:
- ... is some defined normal continuous random variable, usually represented by a capital (uppercase) letter
- N is the arbitrary designation of the normal curve
- μ is the population mean parameter
- σ is the population standard deviation parameter

6.7.2 Probability Density Function for a Normal Random Continuous Variable

This is a specific extension of the idea in Probability Density Function for Random Continuous Variables, but applied specifically to the normal case. Not really used directly for hand calculations due to its complexity.

$$f(...) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-0.50 \cdot (\frac{... - \mu}{\sigma})^2}$$

Where:
- ... is some vector of possible outcomes for an exponential continuous random variable, usually represented as a lowercase letter
- μ is the population mean parameter
- σ is the population standard deviation parameter
- $\cdot$ is an indicator of multiplication
- e is the scientific constant = 2.71828

6.7.3 Notation for a Standard Normal Distribution

This notation is used to describe the core characteristics of a standard normal distributed continuous variable, one composed of z-scores.

$$Z \sim N(0, 1)$$

Where:

- Z is some defined normal continuous random variable transformed into z-scores, usually represented by a capital (uppercase) letter
- N is the arbitrary designation of the normal curve
- μ is the population mean parameter
- σ is the population standard deviation parameter

6.8 Module 7

6.8.1 Formula for the Normal Distribution of Sample Means (CLT)

Much like a Notation for a Normal Distribution, but used specifically in demonstrating central limit theorem (CLT) for means.

$$\bar{\ldots} \sim N(\mu_{\ldots}, \frac{\sigma_{\ldots}}{\sqrt{n}})$$

Where:
- $\ldots$ is some defined continuous random variable, usually represented by a capital (uppercase) letter
- $\bar{\ldots}$ is a random continuous variable consisting of many sample means of variable $\ldots$
- $\mu_{\ldots}$ is the mean of random continuous variable $\ldots$
- $\sigma_{\ldots}$ is the standard deviation of random continuous variable $\ldots$
- n is the size of the individual samples taken to from the sampling distribution (this is consistent across all theoretical samples)

6.8.2 Standard Error of the Mean

The standard error can be thought of as the standard deviation of the sampling distribution of sample means. Put another way, it is hard spread out our mean estimates are in the sampling distribution.

$$SE_{\mu} = \frac{\sigma_{\ldots}}{\sqrt{n}}$$

Where:
- $\ldots$ is some defined continuous random variable, usually represented by a capital (uppercase) letter
- $\sigma_{\ldots}$ is the standard deviation of random continuous variable $\ldots$
- n is the size of the individual samples taken to from the sampling distribution (this is consistent across all theoretical samples)

6.8.3 Z-scores for Sampling Distributions of Sample Means

Much like a normal z-score, but applied to the concept of where a single mean is in the broader sampling distribution.

$$z = \frac{\bar{\dots_1} - \mu_{\dots_2}}{(\frac{\sigma_{\dots_2}}{\sqrt{n}})}$$

Where:
- $\bar{\dots_1}$ is the mean for a single sample of variable $\dots_2$
- $\dots_2$ is some defined continuous random variable, usually represented by a capital (uppercase) letter
- $\mu_{\dots_2}$ is the mean of both $\dots_2$ and $\bar{\dots_1}$
- $\sigma_{\dots_2}$ is the standard deviation of random continuous variable $\dots_2$

6.8.4 Formula for the Normal Distribution of Sample Sums (CLT)

Much like Formula for the Normal Distribution of Sample Means (CLT) but less used in my opinion.

$$\sum \dots N((n)(\mu_{\dots}), (\sqrt{n})(\sigma_{\dots}))$$

Where:
- … is some defined continuous random variable, usually represented by a capital (uppercase) letter
- $\mu_{\dots}$ is the mean of random continuous variable …
- $\sigma_{\dots}$ is the standard deviation of random continuous variable …
- n is the size of the individual samples taken to from the sampling distribution (this is consistent across all theoretical samples)

6.8.5 Z-scores for Sampling Distributions of Sample Sums

See above.

$$z = \frac{\sum \dots_1 - (n)(\mu_{\dots_2})}{(\sqrt(n)(\sigma_{\dots_2}))}$$

Where:
- $\sum \dots_1$ is the sum for a single sample of variable $\dots_2$

- $\ldots_2$ is some defined continuous random variable, usually represented by a capital (uppercase) letter
- $\mu_{\ldots_2}$ is the mean of both $\ldots_2$ and $\bar{\ldots}_1$
- $\sigma_{\ldots_2}$ is the standard deviation of random continuous variable $\ldots_2$