To Do:

Read Chapter 3.

Do end-of-chapter problems 1-7.

Today's and Friday's Lecture

Planning and Conducting Empirical Studies

PPDAC

Designing an Empirical Study – Chapter 3

An empirical study in which data are collected can be used to learn about a population or process.

Such a study can be designed using an algorithm called PPDAC.

PPDAC emphasize the statistical aspects of designing an empirical study.

The Steps of PPDAC

Problem: a clear statement of the study's objectives.

Plan: the procedures used to carry out the study including how the data will be collected.

Data: the physical collection of the data, as described in the Plan.

Analysis: the analysis of the data collected in light of the Problem and the Plan.

Conclusion: The conclusions that are drawn about the Problem and their limitations.

Example: Vitamin D and Influenza A

We use the following research study to illustrate the steps of PPDAC.

***Randomized trial of vitamin D supplementation to prevent seasonal influenza A in schoolchildren*, American Journal of Clinical Nutrition, May 2010, Vol. 91, No. 5, 1255-1260 (http://ajcn.nutrition.org/content/91/5/1255.full.pdf+html)**

The information about the research study on the following slides is taken directly from this article. Articles in medical journals are usually required to follow a prescribed form (similar to PPDAC!).

Example: Vitamin D and Influenza A

Objective:

To investigate the effect of vitamin D supplements on the incidence of seasonal influenza A in schoolchildren.

Study Design:

A randomized, double-blind, placebo-controlled trial was conducted comparing vitamin D supplements (1200 IU/d) with placebo in schoolchildren. The study was conducted by 12 hospitals and 8 doctors in private practice in Japan from December 1, 2008 to March 31, 2009.

Example: Vitamin D and Influenza A

Study Population:

Schoolchildren aged 6-15 were asked to participate in the study by the pediatricians in charge of outpatient clinics.

Children were excluded if they were receiving certain treatments, had allergies to the pills or were unable to swallow pills.

Parents and children were asked to provide written informed consent after the pediatrician explained the study to them at the outpatient clinic.

Example: Vitamin D and Influenza A

Randomization, Blinding:

Using a computer randomization program children were randomly assigned to receive vitamin D or the placebo.

The tablets containing vitamin D or the placebo were identical in appearance.

Only staff at the data monitoring center knew via a randomization code which child received which treatment.

Example: Vitamin D and Influenza A

Outcomes:

The primary outcome was influenza A, diagnosed by medical doctors using a rapid influenza diagnostic test (RIDT) with a nose swab specimen.

The test was conducted following the test manufacturer's protocol.

Example: Vitamin D and Influenza A

Results:

Influenza A occurred in 18 of 167 (10.8%) children in the vitamin D group compared with 31 of 167 (18.6%) children in the placebo group.

The reduction in influenza A was more prominent in children who had not been taking other vitamin D supplements and who started nursery school after age 3.

Example: Vitamin D and Influenza A

Adherence:

Of the 430 children, 334 were followed until the end of the study. Loss to follow-up occurred for 50 children in the vitamin D group and 46 in the placebo group.

Example: Vitamin D and Influenza A

Conclusions:

This study suggests that vitamin D supplementation during the winter may reduce the incidence of influenza A, especially in specific subgroups of schoolchildren.

Steps of PPDAC - The Problem

The elements of the Problem step address questions starting with "What".

To what group of things or people do we want the conclusions to apply (target population)?

What variates can we define?

What is (are) the question(s) we are trying to answer?

What conclusions are we trying to draw?

Steps of PPDAC - The Problem

In the Problem step the units and the target population or target process must be defined.

The Problem - Target Population or Target Process

Definition:

The target population or target process is the collection of units to which the experimenters conducting the empirical study wish the conclusions to apply.

Example: Vitamin D and Influenza A

What is the target population in the Vitamin D/Influenza A study?

Example: Vitamin D and Influenza A

In the vitamin D/influenza A study the units were children.

The target population is all children aged 6-15 in Japan (in the world?) at the time the study was conducted.

Example: Can-Filling Study

In Example 1.5.2 of the Course Notes, two can-filling machines were being compared.

The units were the individual cans.

The target process is all such cans filled now and in the future under current operating conditions.

The Problem - Variates

Definition:

A variate is a characteristic of every unit.

What are the variates of interest in the vitamin D/influenza A study?

Example: Vitamin D and Influenza A

In the vitamin D/influenza A study, for each child (unit) in the target population, the variates of primary interest are whether the child received vitamin D or the placebo and whether or not the child contracted influenza A.

Other variates of interest defined for each unit were age, sex, and height.

Example: Can-Filling Study

In Example 1.5.2 of the Course Notes the purpose of the study was to compare the performance of two can-filling machines.

The volume of liquid in each can (unit) and the machine that filled the can are two variates of interest.

The Problem: Attributes

Definition:

An attribute is a function of the variates over a population.

In the Problem step the questions of interest are specified in terms of the attributes of the target population.

Example: Vitamin D and Influenza A

In the vitamin D/influenza A example what attributes of the target population are of most interest?

Example: Vitamin D and Influenza A

The attributes of the target population of most interest in the vitamin D/influenza A example are:

1) the proportion of children in the target population who would contract influenza A if they took vitamin D

2) the proportion of children in the target population who would contract influenza A if they took the placebo.

Example: Can-Filling Study

In Example 1.5.2 of the Course Notes the average (mean) volume and the variability (variance or standard deviation) of the volumes for all cans filled by each machine under current conditions now and into the future (the target process) are the attributes of interest.

The Problem - Types

The types of problems that an empirical study are designed to study can usually be categorized as one of three types:

Descriptive: to determine a particular attribute of the population (e.g. the national unemployment rate).

Causative: to determine the existence or nonexistence of a causal relationship between two variates (e.g. does a new hockey helmet reduce the risk of concussion).

Predictive: to predict the response of a variate for a given unit (e.g. predict e-cigarette weekly sales if the sales tax on them is doubled).

Example: Vitamin D and Influenza A

In the vitamin D/influenza A study what was the problem of interest?

What type of problem is this?

Example: Vitamin D and Influenza A

The problem of interest was whether taking vitamin D would reduce the proportion of school aged children in Japan who contract influenza A.

Example: Vitamin D and Influenza A

The problem of interest was whether taking vitamin D would reduce the proportion of school aged children in Japan who contract influenza A.

This is a causative problem.

Example: Can-Filling Study

In Example 1.5.2 of the Course Notes a problem of interest was whether the standard deviation of volumes of cans filled by the new machine was less than that of the old machine.

The Problem - Types

In Example 1.5.2 of the Course Notes a problem of interest was whether the standard deviation of volumes of cans filled by the new machine was less than that of the old machine.

This is a descriptive problem.

Causative Problems

Causative problems are common and of critical importance.

The Plan and Analysis steps of an empirical study which is conducted to answer a causative problem must be designed carefully.

More on this in Chapter 8.

Steps of PPDAC - The Plan

The purpose of the Plan step is to decide what units are available for study, what units will be examined, and what variates will be collected and how.

The Plan - Study Population or Study Process

Definition:

The study population or study process is the collection of units available to be included in the study.

The Plan - Study Population or Study Process

Often the study population is a strict subset of the target population.

In many surveys, the study population is a list of people defined by their telephone number.

The Plan - Study Population or Study Process

In many medical applications, animals (e.g. mice) may act as study units when the target population consists of people.

The Plan - Study Population or Study Process

In manufacturing, researchers may want to draw conclusions about the future production process but they can only look at units produced in a laboratory in a pilot process.

In this case, the study units are not part of the target population.

Example: Vitamin D and Influenza A

What is the study population in the Vitamin D/Influenza A study?

Example: Vitamin D and Influenza A

The study population is all children aged 6-15 in Japan who attended the outpatient clinics during the time the study was conducted. (The accrual period was November 1 2008 to December 15 2009.)