Louisa Burns

Osteopathic Research Committee

of

The American Academy of Osteopathy

Research Manual

Revised 8/2001

Louisa Burns Osteopathic Research Committee Manual

Table of Contents

Section I

Section II

Overview of The Louisa Burns Osteopathic Research Committee:

Purpose

The purpose of the Louisa Burns Osteopathic Research Committee (LBORC) is to encourage the development and refinement of Osteopathic principles and practices through clinical research.

The LBORC shall focus on clinical research related to osteopathic manipulative medicine, neuro-musculo-skeletal medicine, and the principles/mechanisms that underlie osteopathic practice; including health care outcomes in the practice of osteopathic manipulative medicine.

Functions

The committee functions in the following ways to provide a unique service to the members of the AAO:

1. LBORC develops, fosters, and supports professional research-quality standards for the collection of clinical data related to osteopathic manipulative medicine and its effects on health and disease.

2. LBORC collaborates in its endeavors with other research supporting organizations such as the AOA Bureau of Research, Education Council of Osteopathic Principles, Research Office of Osteopathic Colleges, OPTI's, NIH, NIMH, AHRQ, etc.

3. LBORC facilitates the development of a database for epidemiologic studies related to osteopathic manipulative medicine.

4. LBORC assists AAO members in developing and performing clinical research. The committee responds to member inquiries at all phases of a research project.

5. LBORC actively seeks new research methods to advance the standards for osteopathic clinical research. These standards are rigorously reviewed and disseminated to the AAO members.

6. LBORC solicits and receives applications for clinical research funding. It is the agency for review for the AAO and recommends the funding of proposed projects. Proposals are evaluated on a peer review basis. While many proposals are recommended for funding, some may not meet the criteria for funding. LBORC can provide a personalized consulting service to assist investigators in initial design of research, grant writing, and grant revision. The purpose of this effort is to increase the likelihood for funding of research related to Osteopathic manipulation and the application of Osteopathic principles and philosophy to the practice of medicine.

7. LBORC monitors the progress of AAO funded projects. The committee is available to the investigator at all times for consultation and to ensure that a project reaches completion.

8. LBORC encourages interdisciplinary collaboration. It is ideal for clinical research projects, providing the added perspective of scientists trained in other disciplines - integrating their expertise with clinical decision making.

9. LBORC provides consultants to the research project within the confines of the budget and subject to the approval by the committee chairperson or designee when it is deemed appropriate. Any consultant(s) selected for such projects will be provided a prospective work agreement that defines the task, schedule, and budget.

10. LBORC will represent the AAO on the subject of Osteopathic research.

Composition of the Committee

This Committee, appointed by the President and confirmed by the board of Trustees of the AAO, shall be composed of a Chairperson and as many members of the committee and consultants as are deemed necessary. It shall include student representation from the UAAO, Council of Interns and Residents, and National Undergraduate Fellows Association that will be appointed by the president of the AAO. Candidates will be selected by their representative organizations for presentation to the president for appointment.

Processing of Research Proposals by the LBORC

Upon receipt of a completed application for research support the Committee shall:

1. Review all proposals. Screen by chairperson and two other committee members. When appropriate the chairperson may also seek the input of additional reviewers.
2. Based on the review process the proposal is classified as follows:
    B. Proposal is not in mission. Investigator notified.
    C. Proposal is within mission but needs to be improved.
        a. Consider referral to consultant to help improve proposal.
        b. Revision by principal investigator with or without help of consultant.
        c. Assist in the rewrite of the proposal.
    C. Proposal is within mission, but funding is deferred due to priorities.

    D. Proposal is within mission. Accepted and funded.
        a. Investigator is notified.
        b. Funding made available according to mutually agreed upon schedule.

Funding, Reports and Publication

A. Investigator provides bi-annual progress reports including up-to-date accounting of the project. It is mandatory that brief research reports be presented to the committee one month prior to the bi-annual meetings (AOA and AAO annual national conferences) in writing or by electronic means for review. Requests for action, report of activity, status of project, and any complications must be included.
B. When a project has more than one funding source, the investigator may request in writing that the reporting and accounting timeline required by **the largest funding agency** be utilized. All efforts to reduce redundancy and expense will be considered by the committee.
C. Investigator provides final summary report complete with full accounting and return of unused funds in a timely fashion.
D. Investigator provides written account of publication plans.
E. Investigator provides a copy of published report to the committee.

RESEARCH RELATED PROGRAMS OF THE LOUISA BURNS OSTEOPATHIC RESEARCH COMMITTEE

Research Funding Resources

The funds available for research in any budgetary year are necessarily limited. Funds are made available through several different foundations and grant resources. Some grants are topic specific, as is the Robuck fund for pediatric research, while other resources are open for all areas of OMM/NMM research. The funding available in each year will not be published to the AAO membership except through the budget of the AAO itself. The LBORC will provide the AAO Board of Trustees with a proposed budget prior to the beginning of the AAO budgetary year. The actual level of funding will be at the discretion of the AAO Board of Trustees. A detailed accounting of the expenditures of the LBORC will be provided to the AAO Board of Trustees within two months of the beginning of the next budgetary year. Use of other funding resources in combination with AAO grants is encouraged. Information on major funding sites in the United States is included below (Assistance for Physicians item 5). International resources can be researched through the World Health Organization (www.who.int/home-page/).

Assistance for Physicians to Develop Clinical Research

Assistance from the LBORC is available to the physician on many levels. Meeting the requirements for clinical research requires considerable planning and expertise to meet the demands of institutional review, data collection and analysis, and preparation for publication. In many cases the physician might feel uncertain about his or her ability to meet accepted scientific standards. The LBORC members and consultants can help the clinician meet these standards at each phase of the inquiry. A written request to the LBORC for assistance in developing acceptable publishable clinical studies may be honored in the following ways:

1. Requests for Funding. The Committee will provide reference materials to aid the clinician through the steps of applying for research funding. Reference materials will include:
    A. Mission Statement
    B. Eligibility requirements
    C. Instruction for applications (deadlines, forms, etc.)
    D. Application forms
    E. Processing of proposals (award notification)
    F. Information for grantees (use of funds, equipment, reports).

2. Provide an "Idea" support package that will assist a clinician in formulating a researchable question and developing a method for answering the question.

3. Provide a "consultation" service to assist the clinician in one or more of the following ways:

A. Identify publishable clinical observations (Case, retro- or prospective clincial studies, reports on practice or participation in osteopathic surveys)
B. Identify a researchable question, assist in selection of data collection methods, assist in planning and conducting the study, and assist in reporting or planning publication
C. Identification of whatever resources that might exist that could assist physicians in development of research in private practice. (Experimental studies generally require affiliation with an institution that can provide an Institutional Review Board and monitoring of approved research)
D. Evaluate resources that are available or needed
E. Provide an estimate of patient availability
F. Provide an estimate of time required to complete the study
G. Provide assistance with the writing of the application for funding or identifying more appropriate funding sources
H. Assist in preparing the results for publication.

4. Provide a guide to osteopathic clinical research to include:
   A. Research standards for osteopathic manipulative treatment
   B. Steps in designing osteopathic research
   C. Sample musculoskeletal forms to record somatic dysfunction.

5. Additional sources of funding for clinical and related research can be found at three Internet sites from the National Institutes of Health and the Agency for Healthcare Research and Quality:
   A. NIH (www.nih.gov)
   B. The National Center for Complementary and Alternative Medicine (NCCAM) (http://altmed.od.nih.gov)
   C. Agency for Healthcare Research and Quality (AHRQ) www.ahcpr.gov.

6. Federal funding is also available to mid-career clinicians who are seeking to become clinical researchers. These grants provide monetary support thus relieving the physician from administrative and patient care responsibilities while involved in patient-oriented research and mentorings (www.nih.gov/grants/funding/phs398/phs398.html).

7. Many physicians may not be formally affiliated with a medical school or research-sponsoring hospital. It is recognized by the LBORC that such physicians might nonetheless be interested in pursuing osteopathic research in their clinical settings. In general, any prospective research proposal will require the review of a group that can function as an Institutional Review Board in determining the safety and appropriateness of the research design for human subjects. In this case the prospective researcher can request the LBORC set up an ad hoc IRB and/or refer the proposal to an already-existing IRB at one of the osteopathic institutions.

The Role of LBORC in Osteopathic Clinical Research

Research Considerations, Models, Methods & Philosophy

This guidebook describes the Louisa Burns Clinical Research Committee program and its procedures for encouraging standard reporting and scientific publication of information concerning osteopathic palpatory diagnosis and osteopathic manipulative treatment of somatic dysfunction as it relates to health and illness.

The committee program will provide assistance to physicians enabling them to participate in the mission of the Louisa Burns Osteopathic Research Committee (LBORC) of the American Academy of Osteopathy (AAO). The LBORC's major focus will be to obtain data relating to osteopathic principles and osteopathic manipulative methods and their effects on health, patient benefits, delivery of health care services and costs. All osteopathic physicians including those in private and group practices; as well as those affiliated with osteopathic medical schools and teaching hospitals, are encouraged to consider contributing to this research effort. The limits of support will be determined by the availability of resources. All candidates must adhere to the accepted requirements for human research.

The program provides mechanisms for:
1. Discussion of ideas in the formative stages
2. Converting ideas into researchable questions
4. Developing the questions into feasible projects
5. Providing criteria for quality reporting in medical publications, preferably refereed publications.

Examples of the types of clinical research related to osteopathic principles and manipulative medicine include but are not limited to:
1. Retrospective studies
2. Systematic documentation of clinical observations and outcomes
3. Prospective cohort studies
4. Prospective blinded studies.

Supportive services, such as research consultation, will be available on an ongoing basis in order to promote interchange between experienced and novice researchers when necessary. The LBORC may also attempt to encourage a connection between the researcher and others involved in similar efforts in order the benefit the researchers and the profession. The LBORC may also facilitate the development of research consortia involving multiple institutions and practices interested in researching questions pertaining to osteopathic principles and manipulative medicine.

Research Considerations

The LBORC is a partner with AACOM, AOA Bureau of Research, AODME and ECOP in the development of a national osteopathic data bank into which individual physicians could contribute data aimed at further elucidating the understanding of osteopathic medicine.

Although the emphasis is on supporting osteopathic principles and the use of osteopathic manipulative medicine in health care, protocols aimed at refuting some of the concepts and/or applications of osteopathic principles and methods will also be endorsed. The traditional roles of clinical research conducted by physicians are described according to the physician's level of affiliation (Figure **1: Traditional Models of Physician Research in Practice)**. The National Institute For Osteopathic Research and Education (NIFORE) mission includes development of methods by which a physician at any tier could participate in any level of research.

Physician Researcher Type	Data Collection Method	Expected Contribution to Osteopathic Medicine
Solo or group practice, Un-Affiliated	Clinical Observation or Descriptive Epidemiology	Observation & Recording, Assessment and Treatment, Outcomes and Professional Training
Affiliation with a Teaching Hospital	Participation in Surveys or Registries	Central Data Bank, Long Terms Benefits of Osteopathic Health Care Length of stay
Affiliation with a Medical College	Publication on Practice	Natural History and Role Of Somatic Dysfunction and Its Treatment In Health, Illness, Disease and Recovery; Health Maintenance
Faculty Member or Professional Staff Member	Experimental research	Cost Effectiveness of OMT
	Epidemiology, Individual or Group Clinical Research, Clinical Trials, Demonstrations	Investigation of etiology, Natural History, and Treatment Variables in Somatic Dysfunction
		Role of Somatic Dysfunction in Health and Recovery from Illness

Figure 1: Traditional Models of Physician Research in Practice:
The traditional roles of Physician participation in research are described across rows according to the level of physician affiliation. The National Institute For Osteopathic Research and Education (NIFORE), due to be launched in the Fall of 2001, will make it possible for physicians in any tier to participate at any level. Larger data pools and cross profession collection will broaden the statistical basis for any study and increase the power of the research generated.

Methods
The research may be either clinical or basic science, supporting or refuting some aspect of osteopathic theory or the application of manipulative methods. However, clinical or basic science research having clear implications in the practice environment has a higher priority than basic science mechanism studies for funding by the LBORC while money designated for research remains tight. While research is being conducted, the use of databanks to capture and retain data is strongly encouraged. Databank systems afford the opportunity to study the natural and long-term history of somatic dysfunction and will provide information on the significance of somatic dysfunction in visceral disease.

1. **The Utility of Data Banks**

    A. Creation of data banks provides an opportunity for the collection of descriptive data on osteopathic practice and has value to the osteopathic profession for physician education, training evaluation, outcomes studies, and for the development of practice standards. Confidentiality (access to information) and the right to privacy (identity) of both the patient and physician are mandatory attributes of such a proposed data bank.

    B. Data banks serve the public interest by evaluating patient and community health needs and services, and by providing data for planning health care delivery and reimbursement. The data bank supplies information on practice, while public health vital statistics, incidence of disease, morbidity and mortality provide information on the collective influence of physician and professional services (and other influences).

2. **Natural History of Somatic Dysfunction and its Significance in Visceral Disease**
The presence of somatic dysfunction as well as its resolution through the use of osteopathic palpatory structural diagnosis and treatment requires carefully recorded observations to determine influences on health and recovery from illness. Attention must be directed to well-defined patient characteristics and interventions provided during a patient's care. These data provide an information base that identifies associations between observed phenomena and are a valuable source for speculation on causal relationships. This will stimulate further clinical questions and promote well-articulated hypotheses for successful research.

    A. How do we measure somatic dysfunction - its presence or absence and its severity?

    B. What is its normal frequency of occurrence in healthy individuals?

    C. What is its normal frequency of occurrence in specific diseases?

    D. What is its measured correlation to visceral disease?

    E. Does OMM treatment impact frequency, severity and or duration of a diseased state?

Philosophy of Research

1. Long-term benefits of osteopathic health care philosophically, historically, and in practice have not had a disease orientation, but rather have emphasized the maintenance and restoration of the body's ability to cope with illness and stress. These benefits have not been adequately evaluated to date and will require including measures reflecting the individual's ability to cope with illness as part of routine patient evaluation.

2. Introduction of "experimental" procedures used in clinical research in addition to tested procedures, provide evidence of causal relationships that advance medical knowledge and improved clinical practice.

3. Clinical outcomes research and cost effectiveness must assess patient benefits and public benefits. Variables of health care services to be considered include the measures of severity of disease, variations in patient characteristics (such as age, sex, medications – both licit and illicit), race and ethnicity, and the duration of disease.

4. Additions to health care services should be based upon clinical studies that are designed to provide evidence of patient benefits. These "Quality Assessment" studies can be based upon advances in medical knowledge, technology, or the introduction of services that regain importance. Studies can be conducted in a variety of venues that may require specific regulatory agency approval and monitoring. All research must formally address city, state, and local codes of ethics along with meeting federal agency criteria.

    A. Data can be accumulated through surveys, registries, reports, and publications.
        a. Data may be reported as part of a data collection system.
        b. Data may be extracted from published resources.
        c. Registries may offer a method of identification of exceptional clinical observations, patient responses or information of medical interest.

    B. Physicians associated with teaching hospitals have opportunities to review personal patient outcomes and compare their services with those of physicians with similar practice profiles. Institutional data on practice is as valuable as data obtained from private practice through surveys, registries, reports, and publications.

    C. Physicians in training should be provided with opportunities to participate in surveys and conduct clinical research and clinical trials. Advancement of knowledge is then gained for the physician and the profession.

    D. Institutional Review Board (IRB) approval for all research is required or a waiver from the IRB must be provided. Research on practice is limited to generalized reports on clinical observations or retrospective studies that maintain privacy and confidentiality. This process must be formally monitored. Clearly stated IRB jurisdictional protocols must be followed.

    E. Medical and research ethics require experimental procedures involving human subjects (patients) to be reviewed, acted upon, approved and monitored by peer

review protecting the rights of the subjects. The standards for peer review have been established and require Institutional Review Boards to be formed, approved and monitored for their compliance with the standards. (See NIH guidelines)

Reviewing The Literature

Purpose:

The literature review provides information on the status of a particular subject. A useful sequence for review is: Medical dictionary for definitions; medical texts for general and specialist level of medical information on a subject; surgical texts for generally accepted procedures and information; special references for an authoritative review of currently accepted procedures and status of knowledge; and a literature review of refereed medical and scientific journals to provide information on what is currently being investigated in your field of interest. The general reference review identifies the usual approach to the topic. It should be recognized that the information in published medical texts is generally five to ten years out of date at the time of publication. The special review identifies options. The literature review indicates current effort by peers in the field.

If the literature is from qualified refereed journals answers to the following questions are of value:

1. What was the research question addressed? What was the Hypothesis?
2. What was the research design? (The choice of a recognized research design for your research is an important choice. Use the literature review to help make a wise choice. You need this information to deal with consultants. They may need your input on selecting the design.)
3. What were the criteria for a confirmed diagnosis?
4. What were the criteria for including or excluding subjects from the research?
5. How are the dependent (outcome) and independent (input) variables defined and measured?
6. What measurements were made to quantify the independent from the dependent variables?
7. What statistical analyses were conducted?
8. How were the data reported?
9. Were the conclusions supported?

The author of a research protocol or publication is judged on the comprehension of the literature as much as on the length of the proposal.

Sources and Relevance

1. Sources of literature and information
    A. Medical libraries
    B. Recent publications

C. Bibliographies from other publications
    a. Textbooks
    b. Refereed journals

D. The Internet, particularly:
    a. The National Library of medicine (http://www.nlm.nih.gov/hinfo.html) and
    b. Medline (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi)

E. The Osteopathic Literature Database (coming online in 2001-2002).

2. Determining whether the literature is relevant to your area of investigation
    A. Osteopathic literature

    B. General medical literature

    C. Basic science literature related to your field of study.

2. Statistics review if needed or for new methods.

4. Preparation for public audience
    A. Style manual (see reading list).

Getting Published - The Report

1. Choose the journal in which you will seek publication
    A. It is useful to know early in the process of research design and data collection what type of journal in which you will be seeking to publish your final data.

    B. Who should know about your study? Why would they be interested?
        a. Osteopathic physicians
        b. Generalists in medicine
        c. Specialists:
            i. in OMM and/or NMM.
            ii. in the area of the medical problem being investigated (e.g. Physiatrists, Neurologists, cardiologists, orthopedists, physical therapists, etc. where the issue relates to the effect of OMT on peripheral blood flow, or immunologists where the research involves the influence of musculoskeletal dysfunction on immune competence).
        d. Basic scientists

    C. Is the journal refereed? You will need to know the style and scientific standards of the journal chosen.

Getting Started

Guidelines for Physicians as Principal Investigator in Clinical Osteopathic Research

1. Sources of potential questions that can lead to research hypotheses:
    A. Examine clinical experience.
    B. Implications of past and current theory.

2. Identify the area of investigation. State the purpose of the study.
    A. Management of a particular illness with manipulation.
    B. Improvement of customary medical care by including OMT.
    C. Viscero-somatic/somato-visceral reflexes.
    D. Association of somatic dysfunction with findings.
    E. Symptomatic relief with manipulation.
    F. Improvement in the maintanance of health by specifically osteopathic intervention (prevention).
    G. Improvement of health care delivery by management using a whole-person paradigm rather than an organ-specific or disease-specific paradigm.
    H. Cost effectiveness of inclusion of musculoskeletal management in the treatment of a particular medical problem.

3. What will be your research design? (See guidelines for design).

4. What is the hypothesis?

5. How will the hypothesis be tested?

6. What question(s) will be asked?

7. What data will be collected?
    A. Patient selection
    B. Patient histories and examinations including
        a. Family history
        b. Socio-economic data
        c. Personal habits such as smoking
        d. Complete presenting complaint
        e. Severity of complaint
        f. Duration of complaint prior to the current intervention
        g. Factors providing relief
        h. Factors making the condition worse
        i. Length and type of prior care for the problem.

    C. Define the method of diagnosis – stating its reliability and validity.

8. Describe your statistical analysis. (See reading list)

9. Identify the risks and potential benefits to the patient.

10. Selection of design.
    A. Design must be able to answer the proposed question.
    B. Design must be feasible in the clinical setting.

11. The data – Identify the methods to be utilized.
    A. Diagnostic criteria
        a. Somatic dysfunction
        b. Abnormal laboratory findings
        c. Abnormal radiological findings
        d. History of a diagnosis of disease
        e. Problems with health but seeking dynamic balance of health
        f. Inclusion: criteria for
        g. Exclusion: criteria for

    B. Repeated measurements before and after treatment
        a. Baseline Data
        b. Outcome Data

    C. How will changes in the patient be measured?
        a. Functional?
        b. Objective vs. subjective?
        c. Measurements and data collection Pre and Post treatment
            i. Observation
            ii. Questionnaires
            iii. Medical records
            iv. Specially designed research forms
            v. Examining physician
            vi. Treating physician
            vii. Instrumental
            viii. Specialized assay

    D. Identify the mode(s) of treatment to be used and the method by which this was determined.

    E. Identify record(s) to be used.
        a. Method(s) of recording somatic findings.
            i. Are they reliable?
            ii. Are there any delays in recording the data?
            iii. Conventional and Osteopathic nomenclature (FRT5), asymmetry, restriction of motion, tissue texture, tenderness.
            iv. Use of standard diagrams.

v. Possible use of voice-activated recorder during exam with transcription later.

c. Development of dictionary - descriptive terms to ensure uniform record.

d. Method of standard exam.

12. Development of a standard treatment procedure.

A. When evaluating OMT technique, the research protocol will use only a single osteopathic manipulative method if OMT is part of the treatment protocol.

B. When evaluating OMT in clinical practice, several different osteopathic manipulative methods may be utilized. Each must be specified as much as possible and the criteria for their use clearly delineated.

C. Where more than one clinician is providing treatment, all must agree on the manipulative method(s), "train" to execute the treatment in a technically equivalent manner, and all should possess roughly equivalent skill in application of the OMT technique(s).

D. Inter-rater reliability must be measured.

13. Assessment of results.

A. Are there any delays in development of effects?

B. Assessment of musculoskeletal findings pre/post treatment?

C. Evaluation of response of medical problem to treatment.

14. Analysis of results for population.

15. Publication.

For all of the types of research listed above the following general questions should be asked:

1. What are the risks/benefits to the patient?
2. Is informed consent necessary from the patient? Has informed consent been obtained?
3. Patient selection criteria

A. Diagnosis

B. Severity

C. Stability

D. Medication

E. Life-style

F. Exclusions

a. Structural

b. Functional

c. Other.

4. Willingness of patient to accept proposed treatment protocol

A. For somatic dysfunction: OMT.

B. For medically defined disease.

5. Identify timeline for completion.

A. Proposal and consents.

B. Institutional Review Board (IRB) approval.

C. Task to be completed.

D. Suggested dates of completion.
E. Actual completion.

Additional Guidelines for Research Designs

Questions to be answered

1. Relate to:
   A. diagnosis
   B. treatment
   C. the patient (activities of daily living)
   D. the condition
   E. the environment of the patient
   F. the environment in which care is provided
   G. pre- and post-treatment
   H. health care delivery
   I. patient education.

2. Bottom Line:
   A. Will the answer benefit the patient?
      a. Directly
      b. Indirectly

   B. Is the treatment proposed cost effective?

   For example: When is a particular treatment indicated in the course of treating peripheral vascular obstruction? The criteria for effective treatment would be the amelioration of symptoms and objective signs: claudication, rest pain, neuropathy, Charcot's joints, muscle and tendon atrophy, skin ulcers, atrophic skin and hair changes, gangrene, edema, vascular insufficiency on arteriogram, venogram, and doppler studies. If we ask what is required to obtain relief, to maintain mobility, to prevent infection and ensure wound healing and list treatments from the most conservative to the most radical, when is each treatment effective?

   Elevate limb
   Osteopathic manipulative treatment
   Exercise
   Vasodilators
   Sympathetcomy
   Vascular reconstruction
   Amputation

Refining the Question is the primary issue.

1. The initial question arises from a hunch, an observation, interaction with peers or the literature.

2. Phrasing the question may require a great deal of thought and time!
3. Define the population to be studied.
4. Operationally define the dependent variables.
5. Operationally define the independent variables.
6. Sampling in order to infer results to a larger population.

Research Design Considerations

1. Are the answers to the questions valid, biased, or influenced by chance?
    A. Validity of the answers.
        a. Are the answers valid for the sample studied?
        b. Are the answers valid for the purported population?

    For more on Validity see **Appendix I**, page 46

    B. Bias in the answers
        a. Selection bias. Is the sample large enough to ensure equal distribution of personal and other characteristics in each comparison group? Was the population randomized?
        b. Measurement bias. Is there a systematic dissimilar method of assessing outcomes in different comparison groups? To minimize measurement bias:
            i. utilize blind observations.
            ii. use multiple independent observers.
            iii. use clearly defined outcome criteria.
            iv. minimize the influence of observing and measuring upon the outcome of the measurements.

2. Does the research design organize the observations in a manner that answers the question? Does the design:
    A. ensure valid answers?
    B. minimize bias?
    C. satisfy the requirements for statistical testing?
    D. ensure ethical treatment of subjects?
    E. satisfy constraints of resources (personnel, time, facilities, testing, and costs)

3. Experimental designs should structure the research to ensure that the conclusion can be unambiguously demonstrated.

4. The sample should be large enough to ensure that the characteristics of the population are equally represented in the sample. This facilitates drawing inferences about the population from the research.

Specific Guidelines for Designing Research Protocols

Osteopathic Manipulative Treatment and the Somatic Dysfunction

In osteopathic clinical practice OMT is considered both a mode of treatment for general medical problems and as a specific treatment for musculoskeletal complaints. Thus, somatic dysfunction and its treatment may be an enabling objective where the purpose is to improve the body's inherent capacity to heal itself, or it can be a terminal objective. All osteopathic clinical studies should document the somatic dysfunction, and the change in the somatic dysfunction following OMT. Since the terminal objective is improvement in a patient's problem(s), careful documentation of all abnormal clinical findings before and after treatment allows assessment of OMT efficacy in treating both the original problem when it was not musculoskeletal, and all relevant musculoskeletal dysfunctions. The patient's problem may be musculoskeletal (pain, inability to perform tasks, etc.), visceral (cough, asthma, peptic ulcer), psycological (depression), or general (activities of daily living, fatigue, weakness, frequent infections). The purpose of the research study is to document improvement in the patient problem or health status.

A study that establishes that OMT improves the somatic dysfunction and is associated with improvement in a patient problem provides the strongest evidence to support osteopathic care by linking the somatic dysfunction to the patient problem. One of the primary purposes of osteopathic clinical research is to provide evidence that osteopathic palpatory diagnosis and manipulative treatment contribute to personal or public health benefits. Identifying changes in somatic dysfunction, changes in health status, and degree of association between the two changes will advance knowledge on the role of the somatic system in health and health care. The development of a theoretical basis for osteopathic practice and the conduct of basic and clinical research to test the theories are equally important but non-essential unless the primary purpose is achieved.

Standards

There are at present no standard procedures for making osteopathic palpatory diagnosis and administering osteopathic manipulative treatment other than the standards maintained for graduation requirements at the Colleges of Osteopathic Medicine. NBOME plans to include a standards examination in the national board exam structure beginning 2001. Researchers, without guidelines or standards, will often utilize different procedures. The absence of standardization makes evaluation of results and the comparison of results between investigators more difficult. If other areas of biomedical science are any indication, it could require considerable time before the profession has developed sufficient information to promulgate standards. Until standards are established, some method is needed to ensure a degree of uniformity in describing somatic dysfunction and recording osteopathic manipulative procedures. This section of the Manual discusses potential guidelines for osteopathic research that could provide a basis for comparison of results. Flexibility is needed in research until standards are established. Consequently guidelines should provide for flexibility while ensuring a common basis for describing the methods used in study of somatic dysfunction. Thus the current guidelines are inherently more open than will be future standards are likely to be. Although the profession has not committed to these guidelines as standards per se, unless the

investigator has clearly defined reasons to deviate from these guidelines it is to the advantage of the profession that they be utilized. If the investigator has good reasons to deviate from these guidelines, they should be clearly stated.

Comments: Standards are established by agreement between experts or developed as the result of research. Establishment or acceptance of standards advances science and practice more rapidly because standards ensure communication and provide a benchmark for comparison. Therefore, the details of inter-examiner standardization must be included in the protocol.

Objectives

1. Provide guidelines for researchers whose research involves osteopathic palpatory and musculoskeletal diagnosis and/or osteopathic manipulative treatment.
2. Provide the same guidelines for reviewers who evaluate osteopathic research protocols, reports or publications.
3. Periodically (every one to five years) evaluate the impact of the guidelines on research proposals, contributions made to knowledge from research that follows the guidelines, and trends towards standardization.
4. Revise or abandon the use of guidelines based upon the evaluations.

Suggested Guidelines

1. Osteopathic palpatory examination (somatic dysfunction):
    A. The examination should be described in research protocols and reported in records, reports and publications in terms of:
        a. Tests (specific palpatory procedure)
        b. Criteria for a positive finding (for each test)
        c. Criteria for an assessment of segmental or regional somatic dysfunction from the findings.
        d. If more than one measure is used, specify concordance among measures.
        e. Variations in findings must be stated.
        f. Evaluation of inter- and intra-examiner reliability
        g. Identification when possible as to whether the somatic dysfunction(s) is (are)
            i. primary
            ii. traumatic
            iii. nontraumatic
            iv. secondary
            v. viscerosomatic.
    B. The description should follow the generally accepted grouping of terms in categories (TART or STAR)
        a. Asymmetry of presentation or motion
        b. Restricted range of motion
        c. Tissue texture change

d. Tenderness or Sensitivity
e. Provide information on location and attributes of somatic dysfunction
    i. segmental or global
    ii. acute/chronic
    iii. severity (3 point scale in the presence of the dysfunction is most supportable)
    iv. history of occurrence
    v. past treatment
    vi. effect of past treatment

C. If the research protocol includes a manipulative treatment arm and a nontreatment arm, both groups should be carefully assessed for the existence and persistence of somatic dysfunction. This could contribute significantly to our understanding of the natural history of somatic dysfunction and its linkage to other disease.

D. What about inter-examiner reliability? Past research in this field has shown that palpatory examination can and often does measure multiple parameters. It is necessary that all examiners are measuring the same thing. An effective method to eliminate measuring different parameters by palpatory examination is to provide a training program to teach examiners to examine and record in a standardized fashion, and evaluate their performance. Alternately, a single examiner needs to be assessed for self-consistency and apply the same bias to all cases. When the somatic dysfunction is evaluated before and after OMT using the same test, inter-examiner reliability is less of an issue and comes into play only when the evaluating physician disagrees with the treating physician on the improvement (or lack thereof) of the somatic dysfunction being treated.

E. These procedures provide the researcher with a method for comparing pre and post treatment status and following the course of somatic dysfunction. Another osteopathic physician or knowledgeable person can visualize how somatic dysfunction and changes in somatic dysfunction were identified.

2. Osteopathic manipulative treatment:
    A. When the objective of the investigation includes OMT, the protocol must propose that osteopathic manipulative treatment to specific and/or general areas of somatic dysfunction benefits the patient's health status.

    B. A musculoskeletal examination must be recorded.

    C. The treatment should include descriptions of:
        a. location
        b. the treatment goal, i.e. what change is expected to occur in the diagnosed somatic dysfunction
        c. technique(s) or methods used
        d. response to treatment in terms of changes in palpatory tests, and

e. the time when changes were observed.

D. Criteria for a positive test to be used for pre and post OMT must be described.

E. After the intervention, the patient is reevaluated using the same test(s) that formed the basis for identification of somatic dysfunction. This is recorded in both the patient's medical record and any research documents that may be separate from the medical record.

F. Research on response to manipulative treatment should categorize the degree of somatic dysfunction response and include in the analysis of results: responders versus non-responders.

G. A study control for reducing bias may use a second blinded physician who does not treat, but evaluates the patient twice, before and after OMT.

H. Unless the protocol specifies a certain procedure, the physician should be free to use those procedures best suited to the patient and physician. This is the model used in actual clinical practice. The physician must document the nature of the OMT method. Osteopathic physicians recognize that with some treatment procedures the tissues and the response of tissues to the physician's intervention provide palpatory information, which directs the treatment. This response to palpatory information is automatic rather than a carefully considered voluntary motor action being directed by the cerebral cortex. As such, the treatment may not be described in precise biomechanical terms. When the OMT protocol is directed to something other than a defined segmental somatic dysfunction (example: total body patterns, fascia, primary respiratory mechanism, etc.), the treatment may be described by the method(s) used (myofascial release, osteopathy in the cranial field [OCF]) rather than a specific procedure. In each case a description of the treatment procedure should be sufficient to allow peer colleagues to understand and reproduce the type of treatment given.

I. Instrumental or objective measurements may be used in research (if available) to support palpatory findings. While the identification of somatic dysfunction is based on palpatory findings (ART, TART {the older anagrams} or STAR), **other objective or instrumental measurements confirming palpatory findings increase the validity of the study if the instrumental measurement is related to a characteristic of somatic dysfunction.).**

J. The protocol must provide for the recording of outcome measurements to document the health benefit of the procedure.

3. The researcher needs to differentiate during development of the research protocol whether the objective is to
   A. Determine the existence of a specific pattern of somatic dysfunction in conjunction with another disease.
      Example: "Cardiac-related somatic dysfunction"

"Musculoskeletal findings in fibromyalgia syndrome"
"Segmental somatic dysfunction differentiating cardiac from esophageal chest pain."

B. Examine the natural history of somatic dysfunction.

C. Determine the effect of OMT on somatic dysfunction and clinical outcomes in another disease.
   a. Studying the response to application of general non-specific articulatory, myofascial or soft tissue techniques.
      Example: "Effect of OMT on serum endorphin levels."
      "Effect of weekly OMT on frequency of upper respiratory infections."
   b. Using protocols that study the effect of a specific manipulative technique or procedure.
      Example: "Use of thoracic pump in the prevention of post-operative telectasis."
      "Effect of upper thoracic Counterstrain in the ICU on MI patients."

D. Study the effect of a particular procedure or method on somatic dysfunction.
   Example: "Duration of thoracic segmental somatic dysfunction with and without OMT."

E. Compare the effects of different manipulative methods.
   Example: "Treatment of sacroiliac somatic dysfunction using cranial OMT versus counter strain."

F. Compare the effects of OMT to other medical interventions.

G. Focus on manipulative practice, i.e., the physician's decisions on technique, duration, intensity and frequency of treatment.

Comment: *These guidelines emphasize the importance of describing and carefully recording somatic dysfunction as well as any changes in somatic dysfunction that accompany treatment. The researcher needs to recognize that changes may not be synchronistic in all of the primary components as defined as asymmetry, range of motion, tenderness, and tissue texture. As an example, range of motion and change in sensitivity may be detectable immediately, while asymmetry and tissue change are delayed. Further, osteopathic manipulative treatment like other treatment modalities, may provide immediate relief or change, or resolution of the condition may require a series of treatments. The researcher should note when single treatments are used to produce long-term changes. Additionally, if a series of treatments are required to reach the treatment goal the sequence should be clearly documented. Finally, the distinctions between variations that occur within normal ranges and those associated*

with visceral or somatic disease must be documented separately to measure treatment efficacy vs. the natural history of somatic dysfunction.

4. Qualification of the individual(s) providing diagnosis or treatment must be stated explicitly.
    A. The treating physician(s) must possess the requisite skill and experience to effectively and safely administer OMT.

    B. The protocol should identify physician(s) qualifications or level of expertise.

    C. The controls for inter- and intra- physician variability should be described if more than one physician is involved or the study extends over a long period of time.

Comment: *Many studies on physicians have emphasized variability in decision-making. Methods should be designed to reduce that variability, and to assess the influence of the variability on the results.*

Controls used when OMT is the Treatment Variable

One of the most perplexing issues in the development of research protocols involving osteopathic manipulative methods is the development of controls.

1. Ethical considerations inevitably arise when controls involve the withholding of manipulative treatment. Both the osteopathic physician and those patients seeking the care of osteopathic physicians are generally convinced of the clinical efficacy of OMT. To withhold a form of treatment in the name of experimental protocols appears to be at odds with the charge of the physician to provide the best care possible. This ethical delimma is not unique to research into OMT. Virtually all clinical research involving a treatment and a non-treatment arm has the same problem. However, this problem can be partly alleviated if the patient is made aware that the protocol is being performed in the name of advancing our knowledge of the efficacy of OMT; that the chance of being included in the treatment and the non-treatment arms is equal and based on random criteria; and that after the study is completed they will have full access to OMT should they wish.

2. Placebo Treatment - It is not possible to effectively administer an osteopathic placebo treatment for the following reasons:
    A. Patients who have experienced OMT know if they are being treated or if they are not.

    B. Treatment of an unrelated area (a frequently suggested form of placebo) does produce change in the body and may produce unanticipated physiologic effects. These effects may be positive or negative.

    C. Patients may drop out of a study because they are not improving as they had expected. (This is a problem in any study.)

D. Placebos can produce statistically significant effects. Is OMT statistically differentiated from placebo and from "standard of care"?

5. Use of non-manipulative interventions as controls. Since the majority of medicine practiced in the Western World does not involve manipulative interventions, comparison of the outcomes of manipulative interventions with and/or without medical intervention is probably ethical. It is acknowledged that some osteopathic physicians do not practice manipulative medicine. However, for those osteopathic physicians who do routinely utilize manipulative treatments in their patient care, ethical considerations as to the withholding of optimum care arise. This is particularly true where the patients themselves accept the same premise. If the osteopathic physician who is performing research belongs to this latter group, the research design may intentionally include only a manipulative arm. In such a case outcomes can be compared to those available in the allopathic literature and/or the known natural history of the disease being studied. In this case it is advisable that the endpoint of the research protocol be an objectively defined variable such as days or blood sugar or come from a measurement instrument already used in the medical literature (e.g. analog pain scales, goniometer readings, three dimensional gait analysis) related to the problem being studied. Purely subjective data, such as assessment of somatic dysfunction pre- and post-intervention by the treating physician can provide clinically relevant information, useful collaborating data, and valid clinical indication of the resolution of somatic dysfunction, but should not be the primary data evaluated for outcomes without ability to compare it to other outcomes measures. The power of the project is greatly influenced by the type of controls utilized, the stability and reliability of the placebo and the accuracy and precision of the measurements made.

Simple Clinical Research Designs: Procedures and Formats

Retrospective Study / Descriptive Study

1. Asking and refining the question.

2. Creating the data collection form.
   Retrospective Study Example: **Sample Form (see Outpatient SOAP note in Appendix III)**
   The Question (study name):

   Patient I.D. BirthDate Sex Height Weight

   Pathology:
   Medical:
   Somatic Dysfunction:
   Care Provided:
   OMT:
   Medication:
   Additional Instructions:
   Outcome:

3. Trial use of form and revision.

4. Complete data collection.

5. Analysis and presentation of data.

6. Development of conclusions concerning the initial question.

7. Publication.

Prospective Clinical Study

1. The question.

2. Use of the retrospective and/or case study as source of questions.
   A. What information was needed that was not available?

   B. What procedures or records were irregular?

C. What inconsistencies are seen among prior studies?

D. What inconsistencies does the researcher see in the results of prior studies as compared to clinical experience?

3. Research design.
   A. Descriptive only.

   B. Observation only.

   C. Comparison with historical controls:
      a. Retrospective study.
      b. Published results in literature.
      c. Network comparison group.

   D. Development of protocol for randomly assigned experimental and control groups.

4. Obtain consent for research.
   A. Departmental and institutional.

   B. Design approval with Human Subjects Committee.

   C. If not being performed in the context of an institution, review and approval by LBORC.

5. Data analysis.

6. Development of conclusions concerning the original hypothesis.

7. Publication.

Comparison Groups - Experimental and Control

1. The refined question.

2. The research design – consider the use a consultant as soon as
   A. the question is refined.
   B. the literature research is completed.
   C. the base line data is defined.
   D. the outcome measure is known.
   E. an estimate of the sample size is known.

3. Obtain departmental and institutional approval (Institutional Review Board).

4. Recruit patients for the study and obtain informed consent.

4. Complete data collection forms.

6. Collect baseline data.

7. Establish controls:
    A. Comparison with historical controls:
        a. Retrospective study.
        b. Published results in literature.
        c. Network comparison group.

    B. Use of the same patient as both research subject and control:
        a. Half of group assigned to control status first then research status.
        b. Half of group assigned to research status first then control status.

    C. Use of simultaneous research and control groups:
        a. Assigned randomly.

        b. Matched populations.

        c. Blinded
            i. physicians / researchers do not know to which group the patient belongs
            ii. patients do not know to which group they belong.

    D. If using a placebo control use a clearly defined one that has been previously proven as neutral in the mechanism(s) being investigated.

8. Prepare tabulations and statistics for presentation.

9. Development of conclusions concerning the original question.

10. Presentation and/or publication.

Criteria for Research Proposals Submitted to the Louisa Burns Osteopathic Research Committee

Projects Requiring Full Committee Review

A. Projects involving human subjects require full review. This information should be submitted in the format described and each item should be addressed. The committee **will only** review proposals that are accompanied by a complete set of documentation.

B. Write a brief abstract summarizing the research to be conducted. This can be identical or similar to the summary required when submitting a proposal to the NIH (200 words or less).

C. Describe the requirements for a subject population. If it is proposed to use special groups such as prisoners, children, the mentally disabled or groups whose ability to give voluntary informed consent may be in question, such concerns must be addressed in detail. Also describe how subjects will be recruited.

4. Analyze the risk/benefit ratios to the patient of the proposed study. As part of this analysis:
   A. Describe and assess any potential risks - physical, psychological, social, legal, economic or other - and assess the likelihood and seriousness of such risks. If the proposed methods of research create potential risks, describe other methods if any, that were considered and why they will not be used.

   B. Describe procedures (including confidentiality safeguards) for protecting against or minimizing potential risks and an assessment of their likely effectiveness. Also, describe procedures for insuring patient anonymity.

   C. Assess the potential benefits to be gained by the individual subjects, as well as benefits, which may accrue to society in general as a result of the planned work.

5. Describe consent procedures to be followed, including how and where informed consent will be obtained.

6. Provide copies of any consent forms to be used. Although written consent forms are usually required, under certain circumstances (e.g., when using questionnaires to collect data) investigators may choose to incorporate the elements of consent into a letter or instruction sheet accompanying the questionnaire.
   A. The consent form, instruction sheet, or explanatory letter should include but need not be restricted to the following statements or concepts:
      a. A reasonable explanation of the research, its purposes, procedures and duration of participation is necessary.

b. A statement to the effect that the experiment has been explained to the subjects and that the subjects understand it, including any inherent risks is needed.
c. The subjects freely consent to participate.
d. The subjects are free to discontinue the experiment at any time without recrimination.
e. All results will be treated with strict confidence and the subjects will remain anonymous.
f. On written request and within the limits of the above restrictions, results will be made available to subjects after the completion of the study.
g. If a treatment is involved, no beneficial effects are guaranteed.
h. When appropriate, the procedure for debriefing the subjects should be stated.
i. If there is a risk of injury to the subject(s), a statement similar to the following must appear in the consent form:

"I understand that in the unlikely event of injury resulting from research procedures, "Some name" State University, its agents, and employees will assume that responsibility as required by law. Emergency medical treatment for injuries or illness is available where the injury or illness is incurred in the course of an experiment. I have been advised that I should look toward my own health insurance program for payment of said medical expenses."

j. Instructions must also be provided regarding who may be contacted for answers to pertinent questions.
k. If the subject is a minor, provisions should be made for obtaining parental or guardian signatures and assurances should be given that the minor's consent will also be obtained.
l. The consent form should not include any exculpatory language whereby the subject waives, or appears to waive, any of his/her legal rights, including any release of the institution or its agents from liability for negligence.
m. A signed copy of such an informed consent form must be on file for each patient included in the research protocol.

7. Submit copies of all gathering instruments (questionnaires, tests, forms, etc.) to be used in the project as part of the application. The method of administering these instruments should also be explained in detail, because the conditions of applying them may be as critical as the instruments themselves.

8. Proposed projects by a graduate student or osteopathic student, resident or fellow submitted to LBORC for review should be accompanied by a signed statement from the sponsoring faculty member(s) stating that he/she has reviewed and approved the proposed project.

9. Provide six copies of the **complete** research proposal in single spaced, 12 point, New Time Roman font or Helvetica font.

10. Proposals may be submitted for review at any time. Because the review process typically requires a minimum of four weeks to complete, investigators should submit the necessary information at least four weeks in advance of the date they wish to initiate their projects. It *is*

strongly recommended that required documents be submitted at least two months prior to the anticipated starting date of the project if at all possible so that unanticipated delays can be minimized.

11. Proposals for prospective research involving the use of human subjects must have been cleared by the human subjects review process at the researchers' osteopathic medical school or hospital prior to submission to the LBORC. Written confirmation of that clearance should be included in the application.

Exemptions from Full Committee Review and Expedited Review

Federal policies allow for two specific modifications in regular review procedures – Exemption and Expedition.

Exemption

1. A researcher may request exemption from full Committee review if the project only involves subjects in one or more of the following categories:
    A. Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as:
        a. research on regular and special education instructional strategies, or
        b. research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

    B. Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), if information taken from these sources is recorded in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

    C. Research involving survey, interview, or observational procedures, **except** where one of the following conditions exist:
        a. Responses are recorded in such a manner that the human subjects can be identified, directly or indirectly through identifiers linked to the subjects which if they became known outside the research, could reasonably place the subjects at risk of criminal or civil liability; or be damaging to the subjects' financial standing or employability.

        b. Responses are recorded in such a manner that the human subjects can be identified, directly or indirectly through identifiers linked to them, and the research deals with drug use, sexual behavior, or the use of alcohol. All research involving survey or interview procedures is exempt without exception, when the respondents are elected or appointed public officials or candidates for public office.

D. Research involving the observation (including observation by participants) of public behavior, except where one of the following conditions exist:
   a. Responses are recorded in such a manner that the human subjects can be identified, directly or indirectly through identifiers linked to the subjects which if they became known outside the research, could reasonably place the subjects at risk of criminal or civil liability; or be damaging to the subjects' financial standing or employability.

   b. Responses are recorded in such a manner that the human subjects can be identified, directly or indirectly through identifiers linked to them, and the research deals with drug use, sexual behavior, or the use of alcohol. All research involving survey or interview procedures is exempt without exception, when the respondents are elected or appointed public officials or candidates for public office.

E. Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

2. Researchers requesting "exemption" from full Committee review should:

   A. Provide one copy of the information required for full Committee review (instead of six copies).

   B. Identify which "exempted" category or categories of the above five categories of research are applicable.

Expedition

1. A researcher may request an **expedited review** of his/her project if it involves no more than minimal risk and only involves human subjects in one or more of the following categories:

   A. Collection of hair and nail clippings, in a non-disfiguring manner, deciduous teeth, and permanent teeth if patient care indicates a need for extraction.

   B. Collection of excreta and external secretions including sweat, uncannulated saliva, placenta removed at delivery, and amniotic fluid at the time of rupture of the membrane prior to or during labor.

   C. Recording of data from subjects 18 years of age or older using non-invasive procedures routinely employed in clinical practice. This includes the use of physical sensors that are applied either to the surface of the body or at a distance and do not involve input of matter or significant amounts of energy into the subject or an invasion of the subject's privacy. It also includes such procedures as weighing, testing sensory acuity, electrocardiography, electroencephalography, thermography, and

detection of electroretinography. It does not include exposure to electromagnetic radiation outside the visible range (for example, x-rays, microwaves).

D. Collection of blood samples by venipuncture, in amounts not exceeding 450 milliliters in an eight-week period and no more often than two times per week, from subjects 18 years of age or older and who are in good health and not pregnant.

E. Collection of both supra- and subgingival dental plaque and calculus, provided the procedure is not more invasive than routine prophylactic scaling of the teeth and the process is accomplished in accordance with accepted prophylactic techniques.

F. Voice recordings made for research purposes such as investigations of speech defects.

G. Moderate exercise by healthy volunteers.

H. The study of existing data, documents, records, pathological specimens, or diagnostic specimens (this applies to most retrospective studies).

I. Research on individual or group behavior or characteristics of individuals, such as studies of perception, cognition, game theory, or test development, where the investigator does not manipulate subjects' behavior and the research will not involve stress to subjects.

J. Research on drugs or devices for which an investigational new drug exemption or an investigational device exemption is not required.

2. Researchers requesting “expedited" review should:

A. Provide two copies of the information required for full Committee review (instead of six copies).

B. Identify which category or categories of the above ten categories of research are applicable.

Conclusion

1. The above statements are considered minimal standards, and in order to meet its obligations the LBORC may need to supplement them in order to adequately protect the rights and welfare of children or others with diminished capacity or requiring special considerations. In determining whether or not such projects dealing with these latter categories are eligible for “exemption” or “expedited” review, special attention will be given to the proposed consent procedures and any risks that obtain primarily because of the vulnerability or dependent nature of the research subjects in these latter categories.
2. It should be pointed out that these two review procedures do not automatically apply to projects involving human research subjects other than competent adults.

3. If it is determined that the project is not eligible for "exemption" or "expedited" review, additional copies of the required information will be requested from the researcher so that the project can be forwarded for full Committee review.

Louisa Burns Osteopathic Research Application for Support

Guidelines for LBORC Applications

Purpose
Osteopathic physicians are being encouraged to publish their creditable observations on effective osteopathic contributions to health care and engage actively in clinical research. A cooperative approach by the committee and physicians will ensure the utility and validity of the insights into osteopathic practice provided by such contributions. It is imperative that the osteopathic community identifies acceptable means of validating the benefits of osteopathic care, thereby providing valuable contributions to medical knowledge. The committee will provide assistance (planning, consulting, financial) for requests that ensure the quality of reported clinician observation.

Guidelines
The committee guidelines suggest factors to be considered in the proposed publication. Attention to these factors will ensure that other physicians will comprehend the presentation. In addition, the guidelines are intended to address the issues that usually decrease the value of published physician observations.

1. An initial statement should indicate what the physician has observed that contributes to patient care. Contributions can include but are not limited to improvements in diagnosis, effects of treatment, criteria for selection of patients most likely to benefit, and clinical observations that identify the influence of examination or treatment.

2. The criteria for a case study, retrospective study of practice or prospective study should be followed. If clinical research is to be used the criteria for the conduct of a clinical research should be used. The common elements of all of the above include:

    A. A detailed description of the patient(s) and condition(s) being reported (or planned for the prospective study).

    B. A detailed description of the osteopathic examination or treatment. The detail should be adequate to allow another physician to repeat the procedure from the written description.

C. A report of the case observation, a factual review of retrospective information from the physician's practice, or the proposed plan for observation to be reported. Sufficient data should be provided to allow the reader to assess the observation and determine whether the information is sufficient to justify the physician's observation of benefit to the patient.

D. Information on how the observation was supported with clinical data or concurrent observations from an unbiased source. Independent observation or measurement is preferable but not always feasible in private practice. This factor receives careful review by the committee.

Ethical Conduct

There is a great deal of concern in the public, academic and professional communities about the ethical professional conduct of researchers, physicians, and authors. Ethical standards are constantly revised to reflect this concern and to provide professional persons with criteria to be used in judging their own actions relative to standard of conduct. The researcher's plans to conduct ethical research should be approved by peer review.

Recipients of AAO-LBORC assistance are responsible for obtaining approval for their research protocol, obtaining informed consent from research subjects and providing protection of the research subject from unexpected risks or invasion of privacy. This approval should be obtained from the College, Hospital, or organization with which they are affiliated. If they are unaffiliated the AAO-LBORC should provide approval based upon our policy.

The care provided by physicians during prospective research should conform to the standards of practice and the procedures in the research protocol. Unusual procedures in the protocol require a risk/benefit ratio evaluation that must be carefully reviewed by qualified physicians. In clinical studies it is expected that the potential benefit to the patient must be significantly higher than the potential or actual risks. The protection of the patient's rights to privacy, e.g., disclosure of information in the medical record, is the responsibility of the researcher and authors.

References for Clinical Research

(Commentary by David P. Yens, Ph.D.)

Bailar, J.C., III, Mosteller, F. (Eds.) Medical Uses of Statistics, 2nd Rd. Boston: NEJM Books, 1992. A good introduction to the topic.

Bakeman, R. Understanding Social Science Statistics, A Spreadsheet Approach. Hillsdale, NJ: Erlbaum, 1992. A nice, fairly easy to use introduction that emphasizes the use of computer spreadsheets.

Corcoran, K. & Fischer, J. Measures for Clinical Practice: Volumes 1 & 2, 3rd Ed. New York: Free Press, 1987. Excellent resource book for measurement instruments.

Creswell, J.W. Research Design: Qualitative & Quantitative Approaches. Thousand Oaks, CA: Sage, 1994. A broader introduction that addresses the new qualitative approach.

Dawson-Sanders, B. & Trapp, R.G. Basic & Clinical Biostatistics. Norwalk, CT: Appleton & Lange, 1994. An excellent, although somewhat dense, introduction to biostatistics; may be a classic.

Dixon R.A., Munro, J.F., & Silcocks, P.B. Evidence Based Medicine Workbook. Butterworth-Heinemann, 1997. A novel workbook introduction to evidence-based medicine concepts and practices.

Fletcher, R.H., Fletcher, S.W., & Wagner, E.H. Clinical Epidemiology: The Essentials. 3rd Ed. Baltimore: Williams & Wilkins, 1996. A fine overview of epidemiology that is organized by specific aspects of health care.

Gehlbach, S.H. Interpreting the Medical Literature, 3rd Ed. New York: McGraw-Hill, 1993. An excellent introduction.

Glantz, S.H. Primer of Biostatistics, 3rd Ed. New York: Mc-Graw Hill, 1992. Another good introduction to the topic.

Hedrick, T.E., Bickman, L., & Rug, D.J. Applied Research Design, A Practical Guide. Thousand Oaks, CA: Sage, 1993. A slim volume that provides a good introduction.

Hulley, S.B. & Cummings, S.R. Designing Clinical Research, an Epidemiological Approach. Baltimore: Williams & Wilkins, 1988. Superb! Probably the best reference at this time.

Ingelfinger, J.A., Mosteller, F., Thibodeau, L.A., & Ware, J.H. Biostatistics in Clinical Medicine, 3rd Ed. New York: McGraw-Hill, 1994. An excellent introduction to biostatistics with strong clinical orientation.

McGaghie, W.C. & Frey, J.J. (Eds.) Handbook for the Academic Physician. New York: Springer-Verlag, 1986. Excellent starting point.

Leedy, P.D. Practical Research: Planning and Design, 5th Ed. New York: Macmillan, 1993. A general introduction, easy to read.

Rosser, W.W. & Shafir, M.S. Evidence-Based Family Medicine. B.C. Decker, 1998. Although oriented toward the family practitioner, this is an excellent description of the process.

Siegel, S. & Castellan, N.J. Nonparametric Statistics for the Behavioral Sciences, 2nd Ed. New York: McGraw-Hill, 2000. Considered the classic in the field; certainly usable for the medical sciences.

Tuckman, B.W. Conducting Educational Research. New York: McGraw-Hill, 1992. Although oriented toward education, this is an excellent introduction to the process of research.

Winer, B.J., Brown, D.R., & Michels, K.M. Statistical Principles in Experimental Design, 3rd Ed. New York: McGraw-Hill, 1991. The classic in the field of statistics.

Yeomans, Steven G. The Clinical Application of Outcomes Assessment. Stamford, CT: Appleton & Lange, 2000. Particularly pertinent to OMT with emphasis on clinical manual medicine outcomes assessments.

Taken From Index Medicus p. 8a7 Sept. 1983
Subject Heading: "Research"

Medical Research: Goals, Problems & Impacts, Castaldi, PA., Pathology, 1983, Jan; 15 (1) 5-10.

Subject Heading: "Research Design"

Challenges of Clinical Measurement, Chalmers, TA., Mt.Sinai J. Med.(NY) 1983, Mar-Apr; 50(2) 138-40.

Taken From Index Medicus Aug. 1983
Subject Heading: "Research"

Special issue on research J. Dent. Educ. 1983 Apr; 47(4} Value 276-8 Planning 258-61 Priority 289-92 Fostering 239-43

Use of Roger's Conceptual System In Research: Kim, HS.,Nursing Review. Mar-Apr; 32(2):89-91.

Taken From Index Medicus 1982 Cumulative
Subject Heading: "Research"

Clin. Investigation On The Threshold of a Golden Era. Pres. Edd. Paul, We; J. Clin Investigation 1981 Sep;68 (3)823-6.

Meeting The Challenge of Research In Fam. Med. Report., Parkerson, GR Jr., J. Fam. Practice 1982 Jan; 14(1):105

Research In Fam Med-Class. Directions & Costs. Perkoff, Gt.; J. Fam. Med. 1981 Sep 13(4):553-7.

Role of Clinician in Biomed. Research. Schiknecht, Hf. Laryngoscope 1982; May 92(5); 487-8.

A Centralized Biomed Res. Data-Process. Unit & Stages of Its Develop. Med Inf (Lond) 1982; Jan-Mar 7(1)39-78.

Form A

Research Project

Do not write in this space
Date received
Amount requested
Rating
Amount granted

Application for Grant from the Louisa Burns Osteopathic Research Committee of the American Academy of Osteopathy

Application is hereby made for a grant in the amount of $ _____ for the period September 1, _____ through August 31, 1_____ for the purpose of conducting a research project on the following subject:

Title of Project:

Name of Principal Investigator:

Title of Principal Investigator:

Institution:

Address:
Co-investigators: Telephone:

Agreement in Regard to Grant As a Result of This Application

The undersigned agree:

(1) to expend funds granted by the American Academy of Osteopathy solely for research purposes specified herein.
(2) to keep careful records of the conduct of this project and keep reasonable care and maintenance on all equipment used.
(3) to return any unexpended funds at the end of the grant period.
(4) to submit four progress reports based on the grant to date. A final report will be submitted four months after the conclusion of the grant period.
(5) that permanent equipment purchased with funds from AAO grants remains the sole property of the AAO and such equipment shall be so designated on the annual financial report.
(6) that when publishing results of investigations, AAO support shall be acknowledged.

(Signed):

Principal Investigator:

Form B

Louisa Burns Osteopathic Research Committee Administrative Data Sheet

Name of principal Investigator:
Title of Project:

Type of Project:

New: Never before submitted to the LBORC, the title must be different from any previous application submitted by the same investigator.

Continuation: Request for support for a study currently, or in the past, supported by the LBORC. The title should be the same as that of the prior application. If the aims of the project have changed significantly, submit the project as a NEW application and note prior grant number:

Resubmission: An application that has been considered by the LBORC in an earlier cycle, but was not funded. The title should be the same as that of the prior application. If the aims of the project have changed significantly, submit the project as a NEW application and note prior grant number:

Other - Specify:

Other Funding:

Please list other agencies that support this type of research. Also list other agencies to which this study has been submitted and any action taken by the agency or the date action is anticipated.

List all other support you currently have for this study and the date funding ceases.

Agency	**Date**

Form C

Louisa Burns Osteopathic Research Committee Certificate of Compliance,

Protection of Research Subjects

Title of project:

Name of principal investigator:

Institution:

I certify to the committee the research protocol, actual application of the design and the facilities in which the research occurs meet the regulations in all federal, state, and local laws concerning the use of human subjects and the handling of experimental animals.***

By: Title: Date:

Committee On Use Of Human Subjects	**Committee On Use Of Animal Welfare**
This project has been reviewed and the following is noted:	This project has been reviewed and the following is noted:
The project does not include activities involving human subjects	The project does not include activities involving animal subjects
The project includes activity involving human products or unidentified patient data, and is exempt from review under DHHS regulations	The project includes activity involving animal products, and exempt from review under DHHS regulations.
Date	Date

Signed:
Title:
Date:

Signed:
Title:
Date:

* DHHS regulation 45 CFR 46 or as revised.

** Regulations in Guide for the Care and Use of Laboratory Animals, and in Principles for Use of Animals, and as established in the Animal Welfare Acts.

Form D

Louisa Burns Osteopathic Research Committee Summary of Budget & Financial Statement

(Attach explanatory sheets as necessary)

Requested Budget

Actual Expenditure

Physical plant renovations

Technician's Salary

Human Research Subjects

Expendable Supplies

Permanent Equipment

Miscellaneous

Totals

Form E

(1 of 2)

Louisa Burns Osteopathic Research Committee Researcher's Personal Data Sheet

Name:
(Last Name) (First Name) (Middle)

Address:

Telephone:

Date of Birth: Place of Birth:

Graduate of which osteopathic college? Year

Postdoctoral education so far (program, institution, years).

Academic degrees other than D.O.

College or University Years Attended Degree

Scientific papers (bibliographic information for published, dates for unpublished)(curriculum vitae may be substituted):

Have you had, or are you applying for, other research or educational grants?

Form E

(2 of 2)

Name: Date: Amount:

Fact Study
Review use of forms - are the forms useful in insuring all data completed, are the forms complete: does data transfer from data collection to data reduction work well?

Research Study
Patient ID, patient demographic data form. Subject files (on clinical research, these files may be coded to assist maintenance of a blind control. Each file has check off form of each record or form generated for that subject. (File the originals, copies for distribution and analysis). Check patient i.d. and research code# before removing personal information.

Analysis Files
These may be computer record or tables. If all data is not tabulated/calculated at once, use a method to identify what subject records are included in tabulation.

Publication
Identify journal and instructions for authors. Date_____

Develop bibliography (Computerized) Date_____

Draft report Date_____

Revise report and prepare illustrations Date_____

Final draft - submit to co-authors and for peer review Date_____

Final report with feedback Date_____

FORM F

Louisa Burns Osteopathic Research Committee Research Project

Management Sheet

Pre-award Activity
LBORC Application - Submitted _____
Approval _____
(File: Application, letter of notification, correspondence)

Detailed Research Plan

	Expected comp. date	Actual comp. date
Schedule: Time line	_____	_____
Research Protocol	_____	_____
Pre-admission Pt. Selection	_____	_____
Diagnosis; Inclusion/Exclusion	_____	_____

File record of: Patient pool - used for inferences from study
Patients identified for study
Patients signing informed consent

Research Subject File
Pre-admission completed checklist

Pre-treatment date collection form (Osteopathic diagnosis)

Data collection- one form or separate form for each session Osteopathic exam, measurements SOAP, OMT

Post treatment data collection forms

Control data (retrospective) or cohort study to identify OMT influence on course of disease

Data Analysis/Reduction Forms
Summary form - tabulation of patient pool,
patient accepted / refused / excluded,
patient dropout

Data tabulation forms
(Try to reduce repeated reference to original records)

Form G

1 of 2

Louisa Burns Osteopathic Research Committee Planning Questionnaire

Name:

Address:

Telephone:

Please answer the below questions:

1. Have you done or been involved in any research in the past?
   If yes, please list.

2. Have you applied as principle investigator from the Louisa Burns Research Committee or from the AOA Bureau of Research in the past for research funding?

3. Have you made a clinical observation or had a case study that led you to the conclusion you wish to prove? What is the question you want to answer?

4. What kind of study are you planning? Retrospective, prospective, case study, review of the literature, other, unknown?

5. What kind of location are you planning to use? Small or large hospital? Office?

6. What background information do you have on this topic already?

7. Have you done a literature search? If so, what sources have you used? Did you find the information that you needed?

Form G
2 of 2

8. What is your background? Are you a clinician, educator, physician, PhD, or some other background? What experience have you had with research?

9. What methods have you considered? Procedures, controls, variables, special techniques, etc.

10. What sources of data collection might be used? Observations, measurements, and/or records, etc.

11. How will you analyze your data? Chi Square, T Test, etc.

12. What specific knowledge will your project produce that relates directly to osteopathic medicine? Examples –
    (1) Osteopathic theory: effects of somatic system on total body performance; interrelations between somatic and other systems; identification for structural, functional, or combinational mechanisms. Osteopathic practice: diagnosis, management, health or disease.

13. Have you completed the NIH on line certification program that allows you to do research on human subjects?

    The website is: (http://ohsr.od.nih.gov/extramural/extramural_training.html) and should take about one and one half hours to complete comfortably.

*These questions are primarily meant to stimulate thought processes. If any of the questions are beyond your initial considerations or if you need help in decisions related to the question, please write “unknown" or “need help". The LBORC will be better equipped to assist in research planning, development, and grant writing with this data. Data, stripped of unique identifiers, will be combined with other applicants’ data for the purpose of tracking the number of new researchers and the growth of the profession in the research environment.

Form H

LBORC Bi-Annual Research Report

Manuscript No.:

Project Title:

Feb Aug

Principal Investigator:

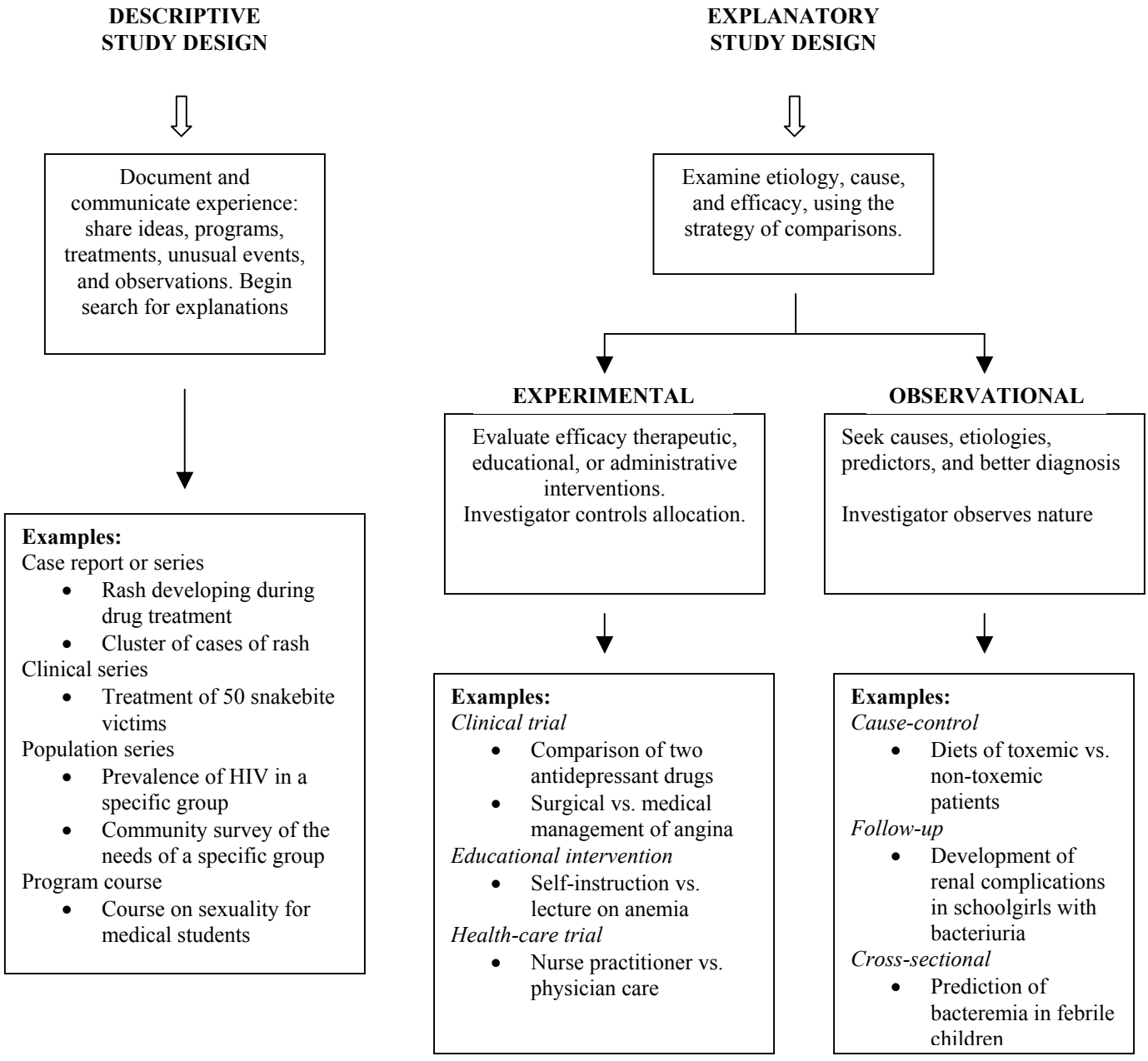
Co-investigators

1. How many patients are currently enrolled in your study?
2. Briefly describe the findings to date.
3. What are the overall results of the study so far?
4. What are the implications of the study so far with respect to osteopathic principles and practice?
5. What is your assessment of the risks versus the benefits based upon study results?
6. Please comment on any new or unusual information obtained·

If informed consent was required by the Institutional Review Board for this study, please attach a copy of the form you are using.

Appendix I

Basic Study Design Flowchart



Modified from: Gehlbach, S.H. Interpreting the Medical Literature, 3rd Edition New York: McGraw-Hill, 1993, p.16

Appendix II
1 of 3

Validity

Factors that invalidate studies:

Internal validity: the data should reflect the characteristics of the sample.

Design problems may affect internal validity. The following list is from Campbell and Stanley "*Experimental and Quasi-experimental Designs for Research*." Chicago: Rand McNally, 1966

1. **History:** External events that may effect the experimental and control groups differently, especially if they are evaluated at different times. Experimental and control groups must participate simultaneously to control for history.

2. **Maturation:** Differential changes that take place in subjects during the study that are not related to the treatment, e.g., children (or the elderly) change rapidly in many ways. If this is a concern, ages of the groups must be equated.

3. **Testing:** Potential effect of a pre-measure or pretest on subsequent evaluations; potential sensitization of the control group to the goals of the study.

4. **Instrumentation:** Changes or instabilities in measurement instruments or research methods during the study; e.g., lack of instrument calibration, taking blood pressures at different sites of the body or at different times.

5. **Statistical Regression:** The tendency for extreme values, outliers, from the mean, to move (regress) toward the mean when measured again. This is one reason why outliers must be evaluated carefully.

6. **Selection:** If groups are not equivalent by age, gender, race and other variables, any differences in the dependent variable are not valid. This is why randomization is required for a valid study.

7. **Experimental Mortality:** If different percentages or numbers of subjects in the groups being compared are lost, the resulting groups may not be comparable. Reasons for this differential loss should be determined.

8. **Stability:** If the findings of the study are unreliable due to variations in the way measures are taken or unreliable instruments, the study results are not valid.

9. **Expectancy:** The expectancy of the experimenter or the subjects may subtly influence the results of the study. To overcome this problem, double blind studies are preferred.

10. **Interactive Combinations:** Combinations of the above problems may lead to severe validity problems.

11. **Compensatory Equalization:** Administration of some treatment to a control group to compensate for its lack of beneficial treatment being received by the experimental group. If something special is provided to the control group to compensate for lack of the treatment, it may affect the outcome by reducing the difference between the groups on the outcome measure.

12. **Compensatory rivalry:** Behavior of a control group such that participants attempt to exceed performance for the experimental group because they are not receiving equal treatment.

13. **Resentful demoralization:** Lowered performance level by the control group because participants resent the lack of experimental treatment.

14. **Diffusion of treatment:** Unintentional administration of treatment to a control group that reduces the post-treatment differences.

Construct Validity

1. **Inadequate preoperational definitions:** If the initial definitions of the variables, treatments, and methods are not adequate, flexibility in the conduct of the study will reduce its validity.

2. **Hypothesis guessing:** Subjects may guess the experimental hypothesis or how the experimenters expect them to behave, and their behavior may reflect the result of their guess. This is related to Expectancy.

3. **Interaction of different treatments:** If more than one treatment is being given, they may interact in unexpected ways to affect the outcome. This is a special concern if subjects are self-medicating with over-the-counter drugs.

4. **Interaction of testing & treatment:** In A-B-A-B designs the testing or evaluation component may interact with the treatments in unexpected ways. Also, present sensitization to the treatment will limit generalization of results.

Statistical Conclusion Validity

1. **Low statistical power:** If the power of the statistical analysis is low, a nonparametric test typically has lower power than a parametric test. If the assumptions of a parametric test were violated, the power would be low.

2. **Violated assumptions of statistical tests:** The assumptions of a parametric test are violated, such as unequal variances or non-normal distributions.

3. **Fishing & error rate:** Performing multiple analyses without adjusting the significance level, or using sequential designs to try to find something significant.

4. **Reliability of measures:** Amount of error in the outcome or dependent measures.

5. **Reliability of treatment implementation:** Amount of error in the treatment, for example, different physicians use a specific osteopathic manipulation differently.

6. **Random irrelevancies in the experimental setting:** Amount of error in the experimental setting other than treatment, for example, subjects are treated differently by staff, specific protocols are not followed exactly, etc.

7. **Random heterogeneity of subjects:** Variation in subjects can affect results. Even random assignment can result in differences between the groups, requiring an analysis of group equivalence after the study.

External validity: the data should reflect the characteristics of the population.

1. **Interaction of selection & treatment:** The subjects in the study may not be the same as other locations and may have the idiosyncratic response to the treatments. Example: if only males are used, the results may not be applicable to females.

2. **Interaction of setting & treatment:** Studies conducted in a unique setting, such as a pressure-packed inner city emergency room, may yield results that are not generalized to other settings.

3. **Interaction of history & treatment:** Some external event that impacts the study, such as a new treatment with demonstrated effectiveness or a change in personnel, may interact with the treatment so that it is not generalized to other time periods.

Bias is a systematic source of error.

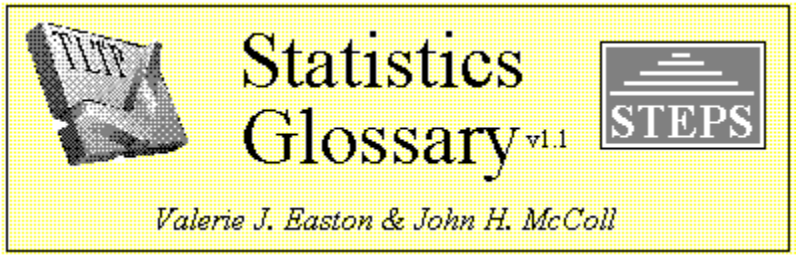
A. Observer (make blind observations) (use multiple independent observers)
B. Sample selection (ensure that comparison groups are as nearly identical as possible)
C. Measurement (calibrate instruments frequently).

Appendix III

Statistics Glossary as adapted from:

Easton, Valerie J., & McColl, John H. Statistics Glossary v1.1

http://www.cas.lancs.ac.uk/glossary_v1.1/hyptest.html#2sampt



can be found at
http://www.cas.lancs.ac.uk/glossary_v1.1/hyptest.html#2sampt

CONTENTS OF GLOSSARY

Probability

Contents

Random Variable
Expected Value
Variance
Probability Distribution
Cumulative Distribution Function
Probability Density Function
Discrete Random Variable
Continuous Random Variable
Probability-Probability (PP) Plot
Quantile-Quantile (QQ) Plot
Normal Distribution
Poisson Distribution
Binomial Distribution
Geometric Distribution
Uniform Distribution
Independent Random Variables
Central Limit Theorem

Outcome

An outcome is the result of an experiment or other situation involving uncertainty.

The set of all possible outcomes of a probability experiment is called a sample space.

Sample Space

The sample space is an exhaustive list of all the possible outcomes of an experiment. Each possible result of such a study is represented by one and only one point in the sample space, which is usually denoted by S.

Examples

1. Experiment Rolling a die once:

Sample space S = {1,2,3,4,5,6}

2. Experiment Tossing a coin:

Sample space S = {Heads, Tails}

3. Experiment Measuring the height (cms) of a girl on her first day at school:

Sample space S = the set of all possible real numbers

Event

An event is any collection of outcomes of an experiment.

Formally, any subset of the sample space is an event.

Any event which consists of a single outcome in the sample space is called an elementary or simple event. Events which consist of more than one outcome are called compound events.

Set theory is used to represent relationships among events. In general, if A and B are two events in the sample space S, then:

$A \cup B$ (A union B) = 'either A or B occurs or both occur'

$A \cap B$ (A intersection B) = 'both A and B occur'

$A \subseteq B$ (A is a subset of B) = 'if A occurs, so does B'

A' or $\bar{A}$ = 'event A does not occur'

ϕ (the empty set) = an impossible event

S (the sample space) = an event that is certain to occur

Example

Experiment Rolling a dice once

Sample space S = {1,2,3,4,5,6}

Events A = 'score < 4' = {1,2,3}

B = 'score is even' = {2,4,6}

C = 'score is 7' = ϕ

$A \cup B$ = 'the score is < 4 or even or both' = {1,2,3,4,6}

$A \cap B$ = 'the score is < 4 and even' = {2}

A' or $\bar{A}$ = 'event A does not occur' = {4,5,6}

Relative Frequency

Relative frequency is another term for proportion; it is the value calculated by dividing the number of times an event occurs by the total number of times an experiment is carried out. The probability of an event can be thought of as its long-run relative frequency when the experiment is carried out many times.

If an experiment is repeated n times, and event E occurs r times, then the relative frequency of the event E is defined to be:

$$rf_n(E) = \frac{r}{n}$$

Example

Experiment Tossing a fair coin 50 times n = 50

Event E = 'heads'

Result 30 heads, 20 tails r = 30

Relative frequency:

$$\text{rfn}(E) = \frac{r}{n} = \frac{30}{50} = \frac{3}{5} = 0.6$$

If an experiment is repeated many, many times without changing the experimental conditions, the relative frequency of any particular event will settle down to some value. The probability of the event can be defined as the limiting value of the relative frequency:

$$P(E) = \lim_{n \to \infty} \text{rfn}(E)$$

For example, in the above experiment, the relative frequency of the event 'heads' will settle down to a value of approximately 0.5 if the experiment is repeated many more times.

Probability

A probability provides a quantitative description of the likely occurrence of a particular event. Probability is conventionally expressed on a scale from 0 to 1; a rare event has a probability close to 0, a very common event has a probability close to 1.

The probability of an event has been defined as its long-run relative frequency. It has also been thought of as a personal degree of belief that a particular event will occur (subjective probability).

In some experiments, all outcomes are equally likely. For example if you were to choose one winner in a raffle from a hat, all raffle ticket holders are equally likely to win, that is, they have the same probability of their ticket being chosen. This is the equally likely outcomes model and is defined to be:

$$P(E) = \frac{\text{number of outcomes corresponding to event E}}{\text{total number of outcomes}}$$

Examples

1. The probability of drawing a spade from a pack of 52 well-shuffled playing cards is:

$$\frac{13}{52} = \frac{1}{4} = 0.25$$

since:

Event E = 'a spade is drawn'
The number of outcomes corresponding to E = 13 (spades) and
The total number of outcomes = 52 (cards)

2. When tossing a coin, we assume that the results 'heads' or 'tails' each have equal probabilities of 0.5.

Subjective Probability

A subjective probability describes an individual's personal judgment about how likely a particular event is to occur. It is not based on any precise computation but is often a reasonable assessment by a knowledgeable person.

Like all probabilities, a subjective probability is conventionally expressed on a scale from 0 to 1; a rare event has a subjective probability close to 0, a very common event has a subjective probability close to 1

A person's subjective probability of an event describes his/her degree of belief in the event.

Example

A Rangers supporter might say 'I believe that Rangers have probability 0.9 of winning the Scottish Premier Division this year since they have been playing really well'.

Independent Events

Two events are independent if the occurrence of one of the events gives us no information about whether or not the other event will occur; that is, the events have no influence on each other.

In probability theory we say that two events, A and B, are independent if the probability that they both occur is equal to the product of the probabilities of the two individual events. That is:

$$P(A \cap B) = P(A).P(B)$$

The idea of independence can be extended to more than two events. For example, A, B and C are independent if:

a) A and B are independent; A and C are independent and B and C are independent (pair wise independence);

b) $P(A \cap B \cap C) = P(A).P(B).P(C)$

If two events are independent then they cannot be mutually exclusive (disjoint) and vice versa.

Example

Suppose that a man and a woman each have a pack of 52 playing cards. Each draws a card from his/her pack. Find the probability that they each draw the ace of clubs.

We define the events:

A = probability that man draws ace of clubs = **1/52**

B = probability that woman draws ace of clubs = **1/52**

Clearly events A and B are independent so:

$$P(A \cap B) = P(A).P(B) = \frac{1}{52}.\frac{1}{52} = 0.00037$$

That is, there is a very small chance that the man and the woman will both draw the ace of clubs.

See also conditional probability.

Mutually Exclusive Events

Two events are mutually exclusive (or disjoint) if it is impossible for them to occur together.

Formally, two events A and B are mutually exclusive if and only if:

$$A \cap B = \phi$$

If two events are mutually exclusive, they cannot be independent and vice versa.

Examples

1. Experiment Rolling a die once

Sample space S = {1,2,3,4,5,6}

Events A = 'observe an odd number' = {1,3,5}

B = 'observe an even number' = {2,4,6}

$A \cap B = \phi$ = the empty set, so A and B are mutually exclusive.

2. A subject in a study cannot be both male and female, nor can they be aged 20 and 30. A subject could however be both male and 20, both female and 30.

Addition Rule

The addition rule is a result used to determine the probability that event A or event B occurs or both occur.

The result is often written as follows, using set notation:

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

where:

P(A) = probability that event A occurs

P(B) = probability that event B occurs

$P(A \cup B)$ = probability that event A or event B occurs

$P(A \cap B)$ = probability that event A and event B occur

For mutually exclusive events, that is events which cannot occur together:

$$P(A \cap B) = 0$$

The addition rule therefore reduces to:

$$P(A \cup B) = P(A) + P(B)$$

For independent events, that is events which have no influence on each other:

$$P(A \cap B) = P(A).P(B)$$

The addition rule therefore reduces to:

$$P(A \cup B) = P(A) + P(B) - P(A).P(B)$$

Example

Suppose we wish to find the probability of drawing either a king or a spade in a single draw from a pack of 52 playing cards.

We define the events A = 'draw a king' and B = 'draw a spade'

Since there are 4 kings in the pack and 13 spades, but 1 card is both a king and a spade, we have:

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52}$$

$$= \frac{16}{52}$$

So, the probability of drawing either a king or a spade is:

$$\frac{16}{52}$$

See also multiplication rule.

Multiplication Rule

The multiplication rule is a result used to determine the probability that two events, A and B, both occur.

The multiplication rule follows from the definition of conditional probability.

The result is often written as follows, using set notation:

$$P(A \cap B) = P(A|B).P(B) \text{ or } P(A \cap B) = P(B|A).P(A)$$

where:

P(A) = probability that event A occurs

P(B) = probability that event B occurs

$P(A \cap B)$ = probability that event A and event B occur

P(A|B) = the conditional probability that event A occurs given that event B has occurred already

P(B|A) = the conditional probability that event B occurs given that event A has occurred already

For independent events, that is events which have no influence on one another, the rule simplifies to:

$$P(A \cap B) = P(A).P(B)$$

That is, the probability of the joint events A and B is equal to the product of the individual probabilities for the two events.

Conditional Probability

In many situations, once more information becomes available, we are able to revise our estimates for the probability of further outcomes or events happening. For example, suppose you go out for lunch at the same place and time every Friday and you are served lunch within 15 minutes with probability 0.9. However, given that you notice that the restaurant is exceptionally busy, the probability of being served lunch within 15 minutes may reduce to 0.7. This is the conditional probability of being served lunch within 15 minutes given that the restaurant is exceptionally busy.

The usual notation for "event A occurs given that event B has occurred" is A|B (A given B). The symbol | is a vertical line and does not imply division. P(A|B) denotes the probability that event A will occur given that event B has occurred already.

A rule that can be used to determine a conditional probability from unconditional probabilities is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

where:

P(A|B) = the (conditional) probability that event A will occur given that event B has occurred already

$P(A \cap B)$ = the (unconditional) probability that event A and event B occur

P(B) = the (unconditional) probability that event B occurs

Law of Total Probability

Bayes' Theorem

Bayes' Theorem is a result that allows new information to be used to update the conditional probability of an event.

Using the multiplication rule:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A).P(A)}{P(B)}$$

This is Bayes' Theorem in its simplest form. Using the Law of Total Probability:

$$P(A|B) = \frac{P(B|A).P(A)}{P(B|A).P(A) + P(B|A').P(A')}$$

where:

P(A) = probability that event A occurs

P(B) = probability that event B occurs

P(A') = probability that event A does not occur

P(A|B) = probability that event A occurs given that event B has occurred already

P(B|A) = probability that event B occurs given that event A has occurred already

P(B|A') = probability that event B occurs given that event A has not occurred already

Random Variable

The outcome of an experiment need not be a number, for example, the outcome when a coin is tossed can be 'heads' or 'tails'. However, we often want to represent outcomes as numbers. A random variable is a function that associates a unique numerical value with every outcome of an experiment. The value of the random variable will vary from trial to trial as the experiment is repeated.

There are two types of random variable - discrete and continuous.

Examples

1. A coin is tossed ten times. The random variable X is the number of tails that are noted. X can only take the values 0,1,....,10, so X is a discrete random variable.

2. A light bulb is burned until it burns out. The random variable Y is its lifetime in hours. Y can take any positive real value, so Y is a continuous random variable.

A random variable has either an associated probability distribution (discrete random variable) or probability density function (continuous random variable).

Expected Value

The expected value (or population mean) of a random variable indicates its average or central value. It is a useful summary value (a number) of the variable's distribution.

Stating the expected value gives a general impression of the behavior of some random variable without giving full details of its probability distribution (if it is discrete) or its probability density function (if it is continuous).

Two random variables with the same expected value can have very different distributions. There are other useful descriptive measures which affect the shape of the distribution, for example variance.

The expected value of a random variable X is symbolized by E(x) or u.

If X is a discrete random variable with possible values x1, x2 ,x3 ,....., xn, and p(xi) denotes P(X = xi), then the expected value of X is defined by:

$$\mu = E(X) = \sum_{\cdot} xi\, p(xi)$$

where the elements are summed over all values of the random variable X.

If X is a continuous random variable with probability density function f(x), then the expected value of X is defined by:

$$\mu = E(X) = \int x\, f(x)\, dx$$

Example

Discrete case : When a die is thrown, each of the possible faces 1,2,3,4,5,6 - the xi's - has a probability of 1/6 - the p(xi)'s - of showing. The expected value of the face showing is therefore:

$$\mu = E(X) = \sum_{\cdot} xi\, p(xi)$$

= (1 x 1/6) +(2 x 1/6) +(3 x 1/6) +(4 x 1/6) +(5 x 1/6) +(6 x 1/6) = 3.5

Notice that, in this case, E(X) is 3.5, which is not a possible value of X.

See also sample mean.

Variance

The (population) variance of a random variable is a non-negative number which gives an idea of how widely spread the values of the random variable are likely to be; the larger the variance, the more scattered the observations on average.

Stating the variance gives an impression of how closely concentrated round the expected value the distribution is; it is a measure of the 'spread' of a distribution about its average value.

Variance is symbolized by V(X) or Var(X) or σ^2

The variance of the random variable X is defined to be:

$$V(X) = \sigma^2 = E\left[X - E(X)\right]^2 = E(X)^2 - \left[E(X)\right]^2$$

where E(X) is the expected value of the random variable X.

Notes

a) the larger the variance, the further that individual values of the random variable (observations) tend to be from the mean, on average;

b) the smaller the variance, the closer that individual values of the random variable (observations) tend to be to the mean, on average;

c) taking the square root of the variance gives the standard deviation. That is:

$$\sqrt{V(X)} = \sqrt{\sigma^2} = \sigma = s$$

d) the variance and standard deviation of a random variable are always non-negative.

See also sample variance.

Probability Distribution

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function.

More formally, the probability distribution of a discrete random variable X is a function which gives the probability p(xi) that the random variable equals xi, for each value xi:

$$p(xi) = P(X = xi)$$

It satisfies the following conditions:

a) $0 \leq p(xi) \leq 1$

b) $\Sigma p(xi) = 1$

Cumulative Distribution Function

All random variables (discrete and continuous) have a cumulative distribution function. It is a function giving the probability that the random variable X is less than or equal to x, for every value x.

Formally, the cumulative distribution function F(x) is defined to be:

$$F(x) = P(X \leq x) \text{ for } -\infty < x < \infty \text{ c.d.f.}$$

For a discrete random variable, the cumulative distribution function is found by summing up the probabilities as in the example below.

For a continuous random variable, the cumulative distribution function is the integral of its probability density function.

Example

Discrete case : Suppose a random variable X has the following probability distribution p(x_i):

xi	0	1	2	3	4	5
p(xi)	1/32	5/32	10/32	10/32	5/32	1/32

This is actually a binomial distribution: $B_i(5,1/2)$ or B(5,1/2). The cumulative distribution function F(x) is then:

xi	0	1	2	3	4	5
F(xi)	1/32	6/32	16/32	26/32	31/32	32/32

F(x) does not change at intermediate values. For example: F (1.3) = F(1.86) = F(1) = 6/32

Probability Density Function

The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval.

More formally, the probability density function, f(x), of a continuous random variable X is the derivative of the cumulative distribution function F(x):

$$f(x) = \frac{d}{dx} F(x)$$

Since $F(x) = P(X \leq x)$ it follows that:

$$\int f(x)\,dx = F(b) - F(a) = P(a < X < b)$$

If f(x) is a probability density function then it must obey two conditions:

a) that the total probability for all possible values of the continuous random variable X is 1:

$$\int f(x)\,dx = 1$$

b) that the probability density function can never be negative: f(x) > 0 for all x.

Discrete Random Variable

A discrete random variable is one which may take on only a countable number of distinct values such as 0,1,2,3,4,..... Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

Compare continuous random variable.

Continuous Random Variable

A continuous random variable is one which takes an infinite number of possible values. Continuous random variables are usually measurements. Examples include height, weight, the amount of sugar in an orange, the time required to run a mile.

Compare discrete random variable.

Probability-Probability (P-P) Plot

A probability-probability (P-P) plot is used to see if a given set of data follows some specified distribution. It should be approximately linear if the specified distribution is the correct model.

The probability-probability (P-P) plot is constructed using the theoretical cumulative distribution function, F(x), of the specified model. The values in the sample of data, in order from smallest to largest, are denoted x(1),x(2),.....,x(n).

For i = 1, 2, ..., n

F(x(i)) is plotted against: $\frac{i - 1/2}{n}$.

Compare Quantile-Quantile (Q-Q) plot.

Quantile-Quantile (QQ) Plot

A Quantile-Quantile (Q-Q) plot is used to see if a given set of data follows some specified distribution. It should be approximately linear if the specified distribution is the correct model.

The Quantile-Quantile (Q-Q) plot is constructed using the theoretical cumulative distribution function, F(x), of the specified model. The values in the sample of data, in order from smallest to largest, are denoted $x_{(1)}$, $x_{(2)}$,....., $x_{(n)}$.

For i = 1, 2,....., n,

$x_{(i)}$ is plotted against: $F^{-1}\left(\frac{i - 1/2}{n}\right)$

Compare probability-probability (P-P) plot.

Normal Distribution

Normal distributions model (some) continuous random variables. Strictly, a Normal random variable should be capable of assuming any value on the real line, though this requirement is often waived in practice. For example, height at a given age for a given gender in a given racial group is adequately described by a Normal random variable even though heights must be positive.

A continuous random variable X, taking all real values in the range $(-\infty, \infty)$, is said to follow a Normal distribution with parameters μ and σ, written:

$$X \sim N(\mu, \sigma^2)$$

if it has probability density function:

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

This probability density function (p.d.f.) is a symmetrical, bell-shaped curve, centered at its expected value μ. The variances σ^2.

Many distributions arising in practice can be approximated by a Normal distribution. Other random variables may be transformed to normality.

The simplest case of the normal distribution, known as the Standard Normal Distribution, has expected value zero and variance one. This is written as N(0,1).

Poisson Distribution

Poisson distributions model (some) discrete random variables. Typically, a Poisson random variable is a count of the number of events that occur in a certain time interval or spatial area. For example, the number of cars passing a fixed point in a 5 minute interval; the number of calls received by a switchboard during a given period of time.

A discrete random variable X is said to follow a Poisson distribution with parameter m, written **X ~ Po(m)**, if it has probability distribution:

$$P(X = x) = \frac{m^x}{x!} e^{-m}$$

where:

x = 0,1,2, ..., n

m > 0.

The following requirements must be met:

a) the length of the observation period is fixed in advance;

b) the events occur at a constant average rate;

c) the number of events occurring in disjoint intervals is statistically independent.

The Poisson distribution has expected value E(X) = m and variance V(X) = m; that is:

$$\mathbf{E(X) = V(X) = m}$$

The Poisson distribution can sometimes be used to approximate the distribution with parameters n and p. When the number of observations n is large, and the success probability p is small, the Bi(n,p) distribution approaches the Poisson distribution with the parameter given by m = np. This is useful since the computations involved in calculating binomial probabilities are greatly reduced.

Binomial Distribution

Binomial distributions model (some) discrete random variables.

Typically, a binomial random variable is the number of successes in a series of trials, for example, the number of 'heads' occurring when a coin is tossed 50 times.

A discrete random variable X is said to follow a Binomial distribution with parameters n and p, written **X ~ Bi(n,p)** or **X ~ B(n,p)**, if it has probability distribution:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where:

x = 0, 1, 2, ..., n

n = 1, 2, 3, ...

p = success probability; 0 < p < 1, and:

$$\binom{n}{x} = \frac{n!}{x!\ (n-x)!}$$

The trials must meet the following requirements:

a) the total number of trials is fixed in advance;

b) there are just two outcomes of each trial; success and failure;

c) the outcomes of all the trials are statistically independent;

d) all the trials have the same probability of success.

The Binomial distribution has expected value E(X) = np and variance V(X) = np (1-p).

Geometric Distribution

Geometric distributions model (some) discrete random variables. Typically, a Geometric random variable is the number of trials required to obtain the first failure, for example, the number of tosses of a coin until the first 'tail' is obtained; components from a production line are tested, in turn, until the first defective item is found.

A discrete random variable X is said to follow a Geometric distribution with parameter p, written **X ~ Ge(p)**, if it has probability distribution:

$$P(X = x) = p^{x-1}\ (1\text{-}p)$$

where:

x = 1, 2, 3, ...n

p = success probability; $0 < p < 1$

The trials must meet the following requirements:

a) the total number of trials is potentially infinite;

b) there are just two outcomes of each trial; success and failure;

c) the outcomes of all the trials are statistically independent;

d) all the trials have the same probability of success.

The Geometric distribution has expected value E(X):

$$E(X) = \frac{1}{1-p}$$

and variance V(X):

$$V(X) = \frac{p}{(1-p)^2}$$

The Geometric distribution is related to the Binomial distribution in that both are based on independent trials in which the probability of success is constant and equal to p. However, a Geometric random variable is the number of trials until the first failure, whereas a Binomial random variable is the number of successes in n trials.

Uniform Distribution

Uniform distributions model (some) continuous random variables and (some) discrete random variables. The values of a uniform random variable are uniformly distributed over an interval. For example, if buses arrive at a given bus stop every 15 minutes, and you arrive at the bus stop at a random time, the time you wait for the next bus to arrive could be described by a uniform distribution over the interval from 0 to 15.

A discrete random variable X is said to follow a Uniform distribution with parameters a and b, written **X ~ Un(a,b)**, if it has probability distribution:

$$P(X = x) = \frac{1}{b - a}$$

where:

x = 1, 2, 3, ..., n.

A discrete uniform distribution has equal probability at each of its n values.

A continuous random variable X is said to follow a Uniform distribution with parameters a and b, written **X ~ Un(a,b)**, if its probability density function is constant within a finite interval [a,b], and zero outside this interval,

$$\text{where } a \leq x \leq b$$

The Uniform distribution has expected value:

$$E(X) = \frac{(a + b)}{2}$$

and variance:

$$V(X) = \frac{1}{12} (b - a)^2$$

Independent Random Variables

Two random variables X and Y say, are said to be independent if and only if the value of X has no influence on the value of Y and vice versa.

The cumulative distribution functions of two independent random variables X and Y are related by

F(x,y) = G(x).H(y)

where:

G(x) and H(y) are the marginal distribution functions of X and Y for all pairs (x,y).

Knowledge of the value of X does not affect the probability distribution of Y and vice versa. Thus there is no relationship between the values of independent random variables.

For continuous independent random variables, their probability density functions are related by

f(x,y) = g(x).h(y)

where:

g(x) and h(y) are the marginal density functions of the random variables X and Y respectively, for all pairs (x,y).

For discrete independent random variables, their probabilities are related by

$P(X = x_i ; Y = y_j) = P(X = x_i).P(Y=y_j)$

for each pair (x_i,y_j).

Central Limit Theorem

Hypothesis Testing

Contents

Hypothesis Test

Setting up and testing hypotheses is an essential part of statistical inference. In order to formulate such a test, usually some theory has been put forward, either because it is believed to be true or because it is to be used as a basis for argument, but has not been proved, for example, claiming that a new drug is better than the current drug for treatment of the same symptoms.

In each problem considered, the question of interest is simplified into two competing claims / hypotheses between which we have a choice; the null hypothesis, denoted H_0, against the alternative hypothesis, denoted H_1. These two competing claims / hypotheses are not however treated on an equal basis, special consideration is given to the null hypothesis. We have two common situations:

1. The experiment has been carried out in an attempt to disprove or reject a particular hypothesis, the null hypothesis, thus we give that one priority so it cannot be rejected unless the evidence

against it is sufficiently strong. For example, H_0: there is no difference in taste between coke and diet coke against H_1: there is a difference.

2. If one of the two hypotheses is 'simpler' we give it priority so that a more 'complicated' theory is not adopted unless there is sufficient evidence against the simpler one. For example, it is 'simpler' to claim that there is no difference in flavor between coke and diet coke than it is to say that there is a difference.

The hypotheses are often statements about population parameters like expected value and variance, for example H_0 might be that the expected value of the height of ten-year-old boys in the Scottish population is not different from that of ten-year-old girls? A hypothesis might also be a statement about the distributional form of a characteristic of interest, for example that the height of ten-year-old boys is normally distributed within the Scottish population?

The outcome of a hypothesis test is 'reject H_0' or 'do not reject H_0'.

Null Hypothesis

The null hypothesis, H_0 represents a theory that has been put forward, either because it is believed to be true or because it is to be used as a basis for argument, but has not been proved. For example, in a clinical trial of a new drug, the null hypothesis might be that the new drug is no better, on average, than the current drug. We would write H_0: there is no difference between the two drugs on average.

We give special consideration to the null hypothesis. This is due to the fact that the null hypothesis relates to the statement being tested, whereas the alternative hypothesis relates to the statement to be accepted if / when the null is rejected.

The final conclusion once the test has been carried out is always given in terms of the null hypothesis. We either 'reject H_0 in favor of H_1' or 'do not reject H_0'; we never conclude 'reject H_1', or even 'accept H_1'.

If we conclude 'do not reject H_0', this does not necessarily mean that the null hypothesis is true, it only suggests that there is not sufficient evidence against H_0 in favor of H_1; rejecting the null hypothesis then, suggests that the alternative hypothesis may be true.

See also hypothesis test.

Alternative Hypothesis

The alternative hypothesis, H_1, is a statement of what a statistical hypothesis test is set up to establish. For example, in a clinical trial of a new drug, the alternative hypothesis might be that the new drug has a different effect, on average, compared to that of the current drug. We would write H_1: the two drugs have different effects, on average. The alternative hypothesis might also be that the new drug is better, on average, than the current drug. In this case we would write H_1: the new drug is better than the current drug, on average.

The final conclusion once the test has been carried out is always given in terms of the null hypothesis. We either 'reject H_0 in favor of H_1' or 'do not reject H_0'; we never conclude 'reject H_1', or even 'accept H_1'.

If we conclude 'do not reject H_0', this does not necessarily mean that the null hypothesis is true, it only suggests that there is not sufficient evidence against H_0 in favor of H_1; rejecting the null hypothesis then, suggests that the alternative hypothesis may be true.

Simple Hypothesis

A simple hypothesis is a hypothesis which specifies the population distribution completely.

Examples

1. H_0: X~Bi(100,1/2) i.e. p is specified

2. H_0: X~N(5,20) i.e. μ and σ^2 are specified

See also composite hypothesis.

Composite Hypothesis

A composite hypothesis is a hypothesis which does not specify the population distribution completely.

Examples

1. X~Bi(100,p) H_1: p > 0.5

2. X~N(0, σ^2) H_1: σ^2 unspecified

See also simple hypothesis.

Type I Error

In a hypothesis test, a type I error occurs when the null hypothesis is rejected when it is in fact true; that is, H_0 is wrongly rejected. For example, in a clinical trial of a new drug, the null hypothesis might be that the new drug is no better, on average, than the current drug; that is H_0: there is no difference between the two drugs on average. A type I error would occur if we concluded that the two drugs produced different effects when in fact there was no difference between them.

The following table gives a summary of possible results of any hypothesis test:

		Decision	
		Reject H_0	**Don't reject H_0**
Truth	**H_0**	Type I Error	Right Decision
	H_1	Right Decision	Type II Error

A type I error is often considered to be more serious, and therefore more important to avoid, than a type II error. The hypothesis test procedure is therefore adjusted so that there is a guaranteed 'low' probability of rejecting the null hypothesis wrongly; this probability is never 0. This probability of a type I error can be precisely computed as,

P(type I error) = significance level = α

The exact probability of a type II error is generally unknown.

If we do not reject the null hypothesis, it may still be false (a type II error) as the sample may not be big enough to identify the falseness of the null hypothesis (especially if the truth is very close to hypothesis).

For any given set of data, type I and type II errors are inversely related; the smaller the risk of one, the higher the risk of the other.

A type I error can also be referred to as an error of the first kind.

Type II Error

In a hypothesis test, a type II error occurs when the null hypothesis H_0, is not rejected when it is in fact false. For example, in a clinical trial of a new drug, the null hypothesis might be that the new drug is no better, on average, than the current drug; that is H_0: there is no difference between the two drugs on average. A type II error would occur if it was concluded that the two drugs produced the same effect, that is, there is no difference between the two drugs on average, when in fact they produced different ones.

A type II error is frequently due to sample sizes being too small.

The probability of a type II error is symbolized by β and written:

P(type II error) = β (but is generally unknown).

A type II error can also be referred to as an error of the second kind.

Compare type I error.
See also power.

Test Statistic

A test statistic is a quantity calculated from our sample of data. Its value is used to decide whether or not the null hypothesis should be rejected in our hypothesis test.

The choice of a test statistic will depend on the assumed probability model and the hypotheses under question.

Critical Value(s)

The critical value(s) for a hypothesis test is a threshold to which the value of the test statistic in a sample is compared to determine whether or not the null hypothesis is rejected.

The critical value for any hypothesis test depends on the significance level at which the test is carried out, and whether the test is one-sided or two-sided.

See also critical region.

Critical Region

The critical region CR, or rejection region RR, is a set of values of the test statistic for which the null hypothesis is rejected in a hypothesis test; that is, the sample space for the test statistic is partitioned into two regions; one region (the critical region) will lead us to reject the null hypothesis H_0', the other not. So, if the observed value of the test statistic is a member of the critical region, we conclude 'reject H_0'; if it is not a member of the critical region then we conclude 'do not reject H_0.

See also critical value.
See also test statistic.

Significance Level

The significance level of a statistical hypothesis test is a fixed probability of wrongly rejecting the null hypothesis H_0, if it is in fact true.

It is the probability of a type I error and is set by the investigator in relation to the consequences of such an error. That is, we want to make the significance level as small as possible in order to protect the null hypothesis and to prevent, as far as possible, the investigator from inadvertently making false claims.

The significance level is usually denoted by α

Significance Level = P(type I error) = α

Usually, the significance level is chosen to be = 0.05 = 5%.

P-Value

The probability value (p-value) of a statistical hypothesis test is the probability of getting a value of the test statistic as extreme as or more extreme than that observed by chance alone, if the null hypothesis H_0, is true.

It is the probability of wrongly rejecting the null hypothesis if it is in fact true.

It is equal to the significance level of the test for which we would only just reject the null hypothesis. The p-value is compared with the significance level and, if it is smaller, the result is significant. That is, if the null hypothesis were to be rejected at

$\alpha = 0.05$, this would be reported as '$p < 0.05$'.

Small p-values suggest that the null hypothesis is unlikely to be true. The smaller it is, the more convincing is the rejection of the null hypothesis. It indicates the strength of evidence for say, rejecting the null hypothesis H_0, rather than simply concluding 'reject H_0' or 'do not reject H_0'.

Power

The power of a statistical hypothesis test measures the test's ability to reject the null hypothesis when it is actually false - that is, to make a correct decision.

In other words, the power of a hypothesis test is the probability of not committing a type II error. It is calculated by subtracting the probability of a type II error from 1, usually expressed as:

Power = 1 - P(type II error) = $(1 - \beta)$

The maximum power a test can have is 1, the minimum is 0. Ideally we want a test to have high power, close to 1.

One-sided Test

A one-sided test is a statistical hypothesis test in which the values for which we can reject the null hypothesis, H_0 are located entirely in one tail of the probability distribution.

In other words, the critical region for a one-sided test is the set of values less than the critical value of the test, or the set of values greater than the critical value of the test.

A one-sided test is also referred to as a one-tailed test of significance.

The choice between a one-sided and a two-sided test is determined by the purpose of the investigation or prior reasons for using a one-sided test.

Example

Suppose we wanted to test a manufacturers claim that there are, on average, 50 matches in a box. We could set up the following hypotheses

$$H_0 : \mu = 50 \text{ against } H_1 : \mu < 50 \text{ or } H_1 : \mu > 50$$

Either of these two alternative hypotheses would lead to a one-sided test. Presumably, we would want to test the null hypothesis against the first alternative hypothesis since it would be useful to know if there is likely to be less than 50 matches, on average, in a box (no one would complain if they get the correct number of matches in a box or more).

Yet another alternative hypothesis could be tested against the same null, leading this time to a two-sided test:

$$H_0 : \mu = 50 \text{ against } H_1 : \mu \neq 50$$

That is, nothing specific can be said about the average number of matches in a box; only that, if we could reject the null hypothesis in our test, we would know that the average number of matches in a box is likely to be less than or greater than 50.

Two-Sided Test

A two-sided test is a statistical hypothesis test in which the values for which we can reject the null hypothesis, H_0 are located in both tails of the probability distribution.

In other words, the critical region for a two-sided test is the set of values less than a first critical value of the test and the set of values greater than a second critical value of the test

A two-sided test is also referred to as a two-tailed test of significance.

The choice between a one-sided test and a two-sided test is determined by the purpose of the investigation or prior reasons for using a one-sided test.

Example

Suppose we wanted to test a manufacturers claim that there are, on average, 50 matches in a box. We could set up the following hypotheses

$$H_0 : \mu = 50 \text{ against } H_1 : \mu < 50 \text{ or } H_1 : \mu > 50$$

Either of these two alternative hypotheses would lead to a one-sided test. Presumably, we would want to test the null against the first alternative hypothesis since it would be useful to know if there is likely to be less than 50 matches, on average, in a box (no one would complain if they get the correct number of matches in a box or more).

Yet another alternative hypothesis could be tested against the same null, leading this time to a two-sided test:

$$H_0 : \mu = 50 \text{ against } H_1 : \mu \neq 50$$

That is, nothing specific can be said about the average number of matches in a box; only that, if we could reject the null hypothesis in our test, we would know that the average number of matches in a box is likely to be less than or greater than 50.

One Sample t-test

A one sample t-test is a hypothesis test for answering questions about the mean where the data are a random sample of independent observations from an underlying normal distribution:

$$N(\mu, \sigma^2), \text{ where } \sigma^2 \text{ is unknown}$$

The null hypothesis for the one sample t-test is:

$$H_0 : \mu = \mu_0 \text{ (where } \mu_0 \text{ known)}$$

That is, the sample has been drawn from a population of given mean and unknown variance (which therefore has to be estimated from the sample).

This null hypothesis, H_0 is tested against one of the following alternative hypotheses, depending on the question posed:

$$H_1 : \mu \neq \mu_0$$
$$H_1 : \mu > \mu_0$$
$$H_1 : \mu < \mu_0$$

Two Sample t-test

A two sample t-test is a hypothesis test for answering questions about the mean where the data are collected from two random samples of independent observations, each from an underlying normal distribution:

$$N(\mu_i, \sigma_i^2), \text{ where } i = 1,2$$

When carrying out a two sample t-test, it is usual to assume that the variances for the two populations are equal, that is:

$$\sigma_1^2 = \sigma_2^2 = \sigma^2$$

The null hypothesis for the two sample t-test is:

$$H_0 : \mu_1 = \mu_2$$

That is, the two samples have both been drawn from the same population.

This null hypothesis is tested against one of the following alternative hypotheses, depending on the question posed.

$$H_1 : \mu_1 \neq \mu_2$$
$$H_1 : \mu_1 > \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

Non-Parametric Methods

Contents

Non-Parametric Tests
Wilcoxon Mann-Whitney Test
Wilcoxon Signed Ranks Test
Sign Test
Runs Test
Kolmogorov-Smirnov Test
Kruskal-Wallis Test

Non-Parametric Tests

Non-Parametric tests are often used in place of their parametric counterparts when certain assumptions about the underlying population are questionable. For example, when comparing two independent samples, the Wilcoxon Mann-Whitney test does not assume that the difference between the samples is normally distributed whereas its parametric counterpart, the two sample t-test does. Non-Parametric tests may be, and often are, more powerful in detecting population differences when certain assumptions are not satisfied.

All tests involving ranked data, i.e. data that can be put in order, are non-parametric.

Wilcoxon Mann-Whitney Test

The Wilcoxon Mann-Whitney Test is one of the most powerful of the non-parametric tests for comparing two populations. It is used to test the null hypothesis that two populations have identical distribution functions against the alternative hypothesis that the two distribution functions differ only with respect to location (median), if at all.

The Wilcoxon Mann-Whitney test does not require the assumption that the differences between the two samples are normally distributed.

In many applications, the Wilcoxon Mann-Whitney Test is used in place of the two sample t-test when the normality assumption is questionable.

This test can also be applied when the observations in a sample of data are ranks, that is, ordinal data rather than direct measurements.

Wilcoxon Signed Ranks Test

The Wilcoxon Signed Ranks test is designed to test a hypothesis about the location (median) of a population distribution. It often involves the use of matched pairs, for example, before and after data, in which case it tests for a median difference of zero.

The Wilcoxon Signed Ranks test does not require the assumption that the population is normally distributed.

In many applications, this test is used in place of the one sample t-test when the normality assumption is questionable. It is a more powerful alternative to the sign test, but does assume that the population probability distribution is symmetric.

This test can also be applied when the observations in a sample of data are ranks, that is, ordinal data rather than direct measurements.

Sign Test

The sign test is designed to test a hypothesis about the location of a population distribution. It is most often used to test the hypothesis about a population median, and often involves the use of matched pairs, for example, before and after data, in which case it tests for a median difference of zero.

The Sign test does not require the assumption that the population is normally distributed.

In many applications, this test is used in place of the one sample t-test when the normality assumption is questionable. It is a less powerful alternative to the Wilcoxon signed ranks test, but does not assume that the population probability distribution is symmetric.

This test can also be applied when the observations in a sample of data are ranks, that is, ordinal data rather than direct measurements.

Runs Test

In studies where measurements are made according to some well-defined ordering, either in time or space, a frequent question is whether or not the average value of the measurement is different at different points in the sequence. The runs test provides a means of testing this.

Example

Suppose that, as part of a screening program for heart disease, men aged 45-65 years have their blood cholesterol level measured on entry to the study. After many months it is noticed that cholesterol levels in this population appear somewhat higher in the Winter than in the Summer. This could be tested formally using a Runs test on the recorded data, first arranging the measurements in the date order in which they were collected.

Kolmogorov-Smirnov Test

For a single sample of data, the Kolmogorov-Smirnov test is used to test whether or not the sample of data is consistent with a specified distribution function. When there are two samples of data, it is used to test whether or not these two samples may reasonably be assumed to come from the same distribution.

The Kolmogorov-Smirnov test does not require the assumption that the population is normally distributed.

Compare Chi-Squared Goodness of Fit Test.

Kruskal-Wallis Test

The Kruskal-Wallis test is a non-parametric test used to compare three or more samples. It is used to test the null hypothesis that all populations have identical distribution functions against

the alternative hypothesis that at least two of the samples differ only with respect to location (median), if at all.

It is the analogue to the F-test used in analysis of variance. While analysis of variance tests depend on the assumption that all populations under comparison are normally distributed, the Kruskal-Wallis test places no such restriction on the comparison.

It is a logical extension of the Wilcoxon-Mann-Whitney Test.

Presenting Data

Contents

Standard Deviation
Coefficient of Variation

Discrete Data

A set of data is said to be discrete if the values / observations belonging to it are distinct and separate, i.e. they can be counted (1,2,3,....). Examples might include the number of kittens in a litter; the number of patients in a doctor's surgery; the number of flaws in one meter of cloth; gender (male, female); blood group (O, A, B, AB).

Compare continuous data.

Categorical Data A set of data is said to be categorical if the values or observations belonging to it can be sorted according to category. Each value is chosen from a set of non-overlapping categories. For example, shoes in a cupboard can be sorted according to color; the characteristic 'color' can have non-overlapping categories 'black', 'brown', 'red' and 'other'. People have the characteristic of 'gender' with categories 'male' and 'female'.

Categories should be chosen carefully since a bad choice can prejudice the outcome of an investigation. Every value should belong to one and only one category, and there should be no doubt as to which one.

Nominal Data

A set of data is said to be nominal if the values / observations belonging to it can be assigned a code in the form of a number where the numbers are simply labels. You can count but not order or measure nominal data. For example, in a data set males could be coded as 0, females as 1; marital status of an individual could be coded as Y if married, N if single.

Ordinal Data

A set of data is said to be ordinal if the values / observations belonging to it can be ranked (put in order) or have a rating scale attached. You can count and order, but not measure, ordinal data.

The categories for an ordinal set of data have a natural order, for example, suppose a group of people were asked to taste varieties of biscuit and classify each biscuit on a rating scale of 1 to 5, representing strongly dislike, dislike, neutral, like, strongly like. A rating of 5 indicates more enjoyment than a rating of 4, for example, so such data are ordinal.

However, the distinction between neighboring points on the scale is not necessarily always the same. For instance, the difference in enjoyment expressed by giving a rating of 2 rather than 1 might be much less than the difference in enjoyment expressed by giving a rating of 4 rather than 3.

Interval Scale

An interval scale is a scale of measurement where the distance between any two adjacent units of measurement (or 'intervals') is the same but the zero point is arbitrary. Scores on an interval scale can be added and subtracted but cannot be meaningfully multiplied or divided. For example, the

time interval between the starts of years 1981 and 1982 is the same as that between 1983 and 1984, namely 365 days. The zero point, year 1 AD, is arbitrary; time did not begin then. Other examples of interval scales include the heights of tides, and the measurement of longitude.

Continuous Data

A set of data is said to be continuous if the values / observations belonging to it may take on any value within a finite or infinite interval. You can count, order and measure continuous data. For example, height; weight; temperature; the amount of sugar in an orange; the time required to run a mile.

Compare discrete data.

Frequency Table

A frequency table is a way of summarizing a set of data. It is a record of how often each value (or set of values) of the variable in question occurs. It may be enhanced by the addition of percentages that fall into each category.

A frequency table is used to summarize categorical, nominal, and ordinal data. It may also be used to summarize continuous data once the data set has been divided up into sensible groups.

When we have more than one categorical variable in our data set, a frequency table is sometimes called a contingency table because the figures found in the rows are contingent upon (dependent upon) those found in the columns.

Example

Suppose that in thirty shots at a target, a marksman makes the following scores:

5 2 2 3 4	4 3 2 0 3	0 3 2 1 5
1 3 1 5 5	2 4 0 0 4	5 4 4 5 5

The frequencies of the different scores can be summarized as:

Score	Frequency	Frequency (%)
0	4	13%
1	3	10%
2	5	17%
3	5	17%
4	6	20%
5	7	23%

Pie Chart

A pie chart is a way of summarizing a set of categorical data. It is a circle which is divided into segments. Each segment represents a particular category. The area of each segment is proportional to the number of cases in that category.

Example

Suppose that, in the last year a sports wear manufacturers has spent 6 million pounds on advertising their products; 3 million has been spent on television adverts, 2 million on sponsorship, 1 million on newspaper adverts, and a half million on posters. This spending can be summarized using a pie chart:

Bar Chart

A bar chart is a way of summarizing a set of categorical data. It is often used in exploratory data analysis to illustrate the major features of the distribution of the data in a convenient form. It displays the data using a number of rectangles, of the same width, each of which represents a particular category. The length (and hence area) of each rectangle is proportional to the number of cases in the category it represents, for example, age group, and religious affiliation.

Bar charts are used to summarize nominal or ordinal data.

Bar charts can be displayed horizontally or vertically and they are usually drawn with a gap between the bars (rectangles), whereas the bars of a histogram are drawn immediately next to each other.

Dot Plot

A dot plot is a way of summarizing data, often used in exploratory data analysis to illustrate the major features of the distribution of the data in a convenient form.

For nominal or ordinal data, a dot plot is similar to a bar chart, with the bars replaced by a series of dots. Each dot represents a fixed number of individuals. For continuous data, the dot plot is similar to a histogram, with the rectangles replaced by dots.

A dot plot can also help detect any unusual observations (outliers), or any gaps in the data set.

Histogram

A histogram is a way of summarizing data that are measured on an interval scale (either discrete or continuous). It is often used in exploratory data analysis to illustrate the major features of the distribution of the data in a convenient form. It divides up the range of possible values in a data set into classes or groups. For each group, a rectangle is constructed with a base length equal to the range of values in that specific group, and an area proportional to the number of observations falling into that group. This means that the rectangles might be drawn of non-uniform height.

The histogram is only appropriate for variables whose values are numerical and measured on an interval scale. It is generally used when dealing with large data sets (>100 observations), when stem and leaf plots become tedious to construct. A histogram can also help detect any unusual observations (outliers), or any gaps in the data set.

Compare bar chart.

Stem and Leaf Plot

A stem and leaf plot is a way of summarizing a set of data measured on an interval scale. It is often used in exploratory data analysis to illustrate the major features of the distribution of the data in a convenient and easily drawn form.

A stem and leaf plot is similar to a histogram but is usually a more informative display for relatively small data sets (<100 data points). It provides a table as well as a picture of the data and from it we can readily write down the data in order of magnitude, which is useful for many statistical procedures.

We can compare more than one data set by the use of multiple stem and leaf plots. By using a back-to-back stem and leaf plot, we are able to compare the same characteristic in two different groups, for example, pulse rate after exercise of smokers and non-smokers.

Box and Whisker Plot (or Boxplot)

A box and whisker plot is a way of summarizing a set of data measured on an interval scale. It is often used in exploratory data analysis. It is a type of graph which is used to show the shape of the distribution, its central value, and variability. The picture produced consists of the most extreme values in the data set (maximum and minimum values), the lower and upper Quartiles, and the median.

A box plot (as it is often called) is especially helpful for indicating whether a distribution is skewed and whether there are any unusual observations (outliers) in the data set.

Box and whisker plots are also very useful when large numbers of observations are involved and when two or more data sets are being compared.

See also 5-Number Summary.

5-Number Summary

A 5-number summary is especially useful when we have so many data that it is sufficient to present a summary of the data rather than the whole data set. It consists of 5 values: the most extreme values in the data set (maximum and minimum values), the lower and upper Quartiles, and the median.

A 5-number summary can be represented in a diagram known as a box and whisker plot. In cases where we have more than one data set to analyze, a 5-number summary is constructed for each, with corresponding multiple box and whisker plots.

Outlier

An outlier is an observation in a data set which is far removed in value from the others in the data set. It is an unusually large or an unusually small value compared to the others.

An outlier might be the result of an error in measurement, in which case it will distort the interpretation of the data, having undue influence on many summary statistics, for example, the mean.

If an outlier is a genuine result, it is important because it might indicate an extreme of behavior of the process under study. For this reason, all outliers must be examined carefully before embarking on any formal analysis. Outliers should not routinely be removed without further justification.

Symmetry

Symmetry is implied when data values are distributed in the same way above and below the middle of the sample.

Symmetrical data sets

a) are easily interpreted;

b) allow a balanced attitude to outliers, that is, those above and below the middle value (median) can be considered by the same criteria;

c) allow comparisons of spread or dispersion with similar data sets.

Many standard statistical techniques are appropriate only for a symmetric distributional form. For this reason, attempts are often made to transform skew-symmetric data so that they become roughly symmetric.

Skewness

Skewness is defined as asymmetry in the distribution of the sample data values. Values on one side of the distribution tend to be further from the 'middle' than values on the other side.

For skewed data, the usual measures of location will give different values, for example, mode<median<mean would indicate positive (or right) skewness.

Positive (or right) skewness is more common than negative (or left) skewness.

If there is evidence of skewness in the data, we can apply transformations, for example, taking logarithms of positive skew data.

Compare symmetry.

Transformation to Normality

If there is evidence of marked non-normality then we may be able to remedy this by applying suitable transformations.

The more commonly used transformations which are appropriate for data which are skewed to the right with increasing strength (positive skew) are:

$\frac{1}{x}$, $\log x$ and $\sqrt{x}$

The more commonly used transformations which are appropriate for data which are skewed to the left with increasing strength (negative skew) are:

x^2, x^3 and e^x

where the x's are the data values.

Scatter Plot

A scatter plot is a useful summary of a set of bivariate data (two variables), usually drawn before working out a linear correlation coefficient or fitting a regression line. It gives a good visual picture of the relationship between the two variables, and aids the interpretation of the correlation coefficient or regression model.

Each unit contributes one point to the scatter plot, on which points are plotted but not joined. The resulting pattern indicates the type and strength of the relationship between the two variables.

ILLUSTRATIONS

a) The more the points tend to cluster around a straight line, the stronger the linear relationship between the two variables (the higher the correlation).

b) If the line around which the points tends to cluster runs from lower left to upper right, the relationship between the two variables is positive (direct).

c) If the line around which the points tends to cluster runs from upper left to lower right, the relationship between the two variables is negative (inverse).

d) If there exists a random scatter of points, there is no relationship between the two variables (very low or zero correlation).

e) Very low or zero correlation could result from a non-linear relationship between the variables. If the relationship is in fact non-linear (points clustering around a curve, not a straight line), the correlation coefficient will not be a good measure of the strength.

A scatter plot will also show up a non-linear relationship between the two variables and whether or not there exist any outliers in the data.

More information can be added to a two-dimensional scatter plot - for example, we might label points with a code to indicate the level of a third variable.

If we are dealing with many variables in a data set, a way of presenting all possible scatter plots of two variables at a time is in a scatter plot matrix. ILLUSTRATION.

Sample Mean

The sample mean is an estimator available for estimating the population mean . It is a measure of location, commonly called the average, often symbolized $\bar{x}$

Example

Lets say our data set is: 5 3 54 93 83 22 17 19.

The sample mean is calculated by taking the sum of all the data values and dividing by the total number of data values:

$$\bar{x} = \frac{5+3+54+93+83+22+17+19}{8} = 37$$

Its value depends equally on all of the data which may include outliers. It may not appear representative of the central region for skewed data sets.

It is especially useful as being representative of the whole sample for use in subsequent calculations.

See also expected value.

Median

The median is the value halfway through the ordered data set, below and above which there lies an equal number of data values.

Example

With an odd number of data values, for example 21, we have:

Data	96 48 27 72 39 70 7 68 99 36 95 4 6 13 34 74 65 42 28 54 69
Ordered Data	4 6 7 13 27 28 34 36 39 42 48 54 65 68 69 70 72 74 95 96 99
Median	48, leaving 10 values below and 10 values above

With an even number of data values, for example 20, we have:

Data	57 55 85 24 33 49 94 2 8 51 71 30 91 6 47 50 65 43 41 7
Ordered Data	2 6 7 8 24 30 33 41 43 47 49 50 51 55 57 65 71 85 91 94
Median	Halfway between the two 'middle' data points - in this case halfway between 47 and 49 = 48

It is generally a good descriptive measure of the location which works well for skewed data, or data with outliers.

The median is the 0.5 Quantile.

Mode

The mode is the most frequently occurring value in a set of discrete data. There can be more than one mode if two or more values are equally common.

Example

Suppose the results of an end of term Statistics exam were distributed as follows:

Student:.....Score:

1.....94

2.....81

3.....56

4.....90

5.....70

6.....65

7.....90

8.....90

9.....30

Then the **mode** (most common score) is 90, and the **median** (middle score) is 81.

Dispersion

The data values in a sample are not all the same. This variation between values is called dispersion.

When the dispersion is large, the values are widely scattered; when it is small they are tightly clustered. The width of diagrams such as dot plots, box plots, stem and leaf plots is greater for samples with more dispersion and vice versa.

There are several measures of dispersion, the most common being the standard deviation. These measures indicate to what degree the individual observations of a data set are dispersed or 'spread out' around their mean.

In manufacturing or measurement, high precision is associated with low dispersion.

Range

The range of a sample (or a data set) is a measure of the spread or the dispersion of the observations. It is the difference between the largest and the smallest observed value of some quantitative characteristic and is very easy to calculate.

A great deal of information is ignored when computing the range since only the largest and the smallest data values are considered; the remaining data are ignored.

The range value of a data set is greatly influenced by the presence of just one unusually large or small value in the sample (outlier).

Examples

1. The range of 65,73,89,56,73,52,47 is 89-47 = 42.

2. If the highest score in a 1st year statistics exam was 98 and the lowest 48, then the range would be 98-48 = 50.

Inter-Quartile Range (IQR)

The inter-Quartile range is a measure of the spread of or dispersion within a data set.

It is calculated by taking the difference between the upper and the lower Quartiles. For example:

Data	2 3 4 5 6 6 6 7 7 8 9
Upper Quartile	7
Lower Quartile	4
IQR	7 - 4 = 3

The IQR is the width of an interval which contains the middle 50% of the sample, so it is smaller than the range and its value is less affected by outliers.

Quantile

Quantiles are a set of 'cut points' that divide a sample of data into groups containing (as far as possible) equal numbers of observations.

Examples of Quantiles include Quartile, Quintile, and Percentile.

Percentile

Percentiles are values that divide a sample of data into one hundred groups containing (as far as possible) equal numbers of observations. For example, 30% of the data values lie below the 30th Percentile.

See Quantile.

Compare: Quintile, Quartile.

Quartile

Quartiles are values that divide a sample of data into four groups containing (as far as possible) equal numbers of observations.

A data set has three Quartiles. References to Quartiles often relate to just the outer two, the upper and the lower Quartiles; the second Quartile being equal to the median. The lower Quartile is the data value a quarter way up through the ordered data set; the upper Quartile is the data value a quarter way down through the ordered data set.

Example

Data	6 47 49 15 43 41 7 39 43 41 36
Ordered Data	6 7 15 36 39 41 41 43 43 47
Median	41
Upper Quartile	43
Lower Quartile	15

See Quantile.
Compare: Percentile, Quintile.

Quintile

Quintiles are values that divide a sample of data into five groups containing (as far as possible) equal numbers of observations.

See Quantile.

Compare Quartile, Percentile.

Sample Variance

Sample variance is a measure of the spread of or dispersion within a set of sample data.

The sample variance is the sum of the squared deviations from their average divided by one less than the number of observations in the data set. For example, for n observations x_1, x_2, x_3,....., x_n with sample mean:

$$\bar{x} = \frac{\sum x_i}{n}$$

the sample variance is:

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

See also variance.

Standard Deviation

Standard deviation is a measure of the spread or dispersion of a set of data.

It is calculated by taking the square root of the variance and is symbolized by s.d, or s. That is:

$$\sqrt{V(X)} = \sqrt{\sigma^2} = s$$

The more widely the values are spread out, the larger the standard deviation. For example, say we have two separate lists of exam results from a class of 30 students; one ranges from 31% to

98%, the other from 82% to 93%, then the standard deviation would be larger for the results of the first exam.

Coefficient of Variation

The coefficient of variation measures the spread of a set of data as a proportion of its mean. It is often expressed as a percentage.

It is the ratio of the sample standard deviation to the sample mean:

$$\frac{s}{\bar{x}}$$

There is an equivalent definition for the coefficient of variation of a population, which is based on the expected value and the standard deviation of a random variable:

$$\frac{\sqrt{V(X)}}{E(X)}$$

Categorical Data

Contents

Contingency Table

A contingency table is a way of summarizing the relationship between variables, each of which can take only a small number of values. It is a table of frequencies classified according to the values of the variables in question.

When a population is classified according to two variables it is said to have been 'cross-classified' or subjected to a two-way classification. Higher classifications are also possible.

A contingency table is used to summarize categorical data. It may be enhanced by including the percentages that fall into each category.

What you find in the rows of a contingency table is contingent upon (dependent upon) what you find in the columns.

Confidence Interval for a Proportion

A confidence interval gives us some idea of the range of values which an unknown population parameter (such as the mean μ, variance σ^2) is likely to take based on a given set of sample data.

Sometimes we are interested in the proportion of responses that fall into one of two categories. For example, a firm may wish to know what proportion of their customers pay by credit card as opposed to those who pay by cash; the manager of a T.V. station may wish to know what percentage of households in a certain town have more than one T.V. set; a doctor may be interested in the proportion of patients who benefited from a new drug as opposed to those who didn't, etc. A confidence interval for a proportion would specify a range of values within which the true population proportion may lie, for such examples.

The procedure for obtaining such an interval is based on the proportion, p of a sample from the overall population.

Confidence Interval for the Difference Between Two Proportions

A confidence interval gives us some idea of the range of values which an unknown population parameter (such as the mean μ, variance σ^2) is likely to take based on a given set of sample data.

Many occasions arise where we have to compare the proportions of two different populations. For example, a firm may want to compare the proportions of defective items produced by different machines; medical researchers may want to compare the proportions of men and women who suffer heart attacks etc. A confidence interval for the difference between two proportions would specify a range of values within which the difference between the two true population proportions may lie, for such examples.

The procedure for obtaining such an interval is based on the sample proportions, p_1 and p_2, from their respective overall populations.

Expected Frequencies

In contingency table problems, the expected frequencies are the frequencies that you would predict ('expect') in each cell of the table, if you knew only the row and column totals, and if you assumed that the variables under comparison were independent.

See also contingency table.

Observed Frequencies

In contingency table problems, the observed frequencies are the frequencies actually obtained in each cell of the table, from our random sample. When conducting a chi-squared test, the term observed frequencies is used to describe the actual data in the contingency table.

Observed frequencies are compared with the expected frequencies and differences between them suggest that the model expressed by the expected frequencies does not describe the data well.

See also contingency table.

Chi-Squared Goodness of Fit Test

The Chi-Squared Goodness of Fit Test is a test for comparing a theoretical distribution, such as a Normal, Poisson etc, with the observed data from a sample.

Chi-Squared Test of Association

The Chi-Squared Test of Association allows the comparison of two attributes in a sample of data to determine if there is any relationship between them.

The idea behind this test is to compare the observed frequencies with the frequencies that would be expected if the null hypothesis of no association / statistical independence were true. By assuming the variables are independent, we can also predict an expected frequency for each cell in the contingency table.

If the value of the test statistic for the chi-squared test of association is too large, it indicates a poor agreement between the observed and expected frequencies and the null hypothesis of independence / no association is rejected.

Chi-Squared Test of Homogeneity

On occasion it might happen that there are several proportions in a sample of data to be tested simultaneously. An even more complex situation arises when the several populations have all been classified according to the same variable. We generally do not expect an equality of proportions for all the classes of all the populations. We do however, quite often need to test whether the proportions for each class are equal across all populations and whether this is true for each class. If this proves to be the case, we say the populations are homogeneous with respect to the variable of classification. The test used for this purpose is the Chi-Squared Test of Homogeneity, with hypotheses,

H_0: the populations are homogeneous with respect to the variable of classification

against

H_1: the populations are not homogeneous