

Statistical Inferential Reasoning

Courses that would qualify to be one of the “3-hour course(s) devoted to a conceptual and practical understanding of statistical inferential reasoning” should be focused on the student’s ability to evaluate the efficacy of claims based on statistical constructs and to understand and articulate important risks that these claims often address, both through the formal science of statistical inference and informal activity of human inference. These courses should not have computations and derivations as their primary focus; neither should they be abstract reasoning courses devoid of numerical data.

Toward that end, it is expected that any course that qualifies must exhibit a syllabus that offers convincing evidence that, upon successfully completing this course, students will be able to:

- A. (At least 25% of the course) - Evaluate common claims arising from the formal statistical inference conveyed in margins of error and confidence intervals. Students must be able to articulate the sense in which margins of error and confidence intervals address and purport to quantify risks that are of practical interest. Although skill in the computation of these quantities is an acceptable by-product, the demonstrated skill set **must not** be confined to, or even largely focused on the computation of these quantities,. In particular, the student must:
 - 1. Be able to connect the uncertainty of sampling variability with margins of error and confidence intervals. This connection needs to be formal in the sense that the student needs to be able to demonstrate an understanding of the roles of sampling distributions, and standard scores, as well as the central limit theorem (non-mathematical treatment) in the production, but more importantly, the interpretation of margins of error and confidence intervals.
 - 2. Be able to demonstrate an understanding that some of the other major sources of uncertainty, such as biased samples and questionnaires that are worded in a biased or misleading fashion are not addressed by margins of error or confidence intervals.
- B. (At least 25% of the course) - Evaluate common claims arising from the formal statistical inference conveyed in null hypothesis testing associated with statistically designed experiments. Students must be able to articulate the sense in which null hypothesis testing addresses and purports to quantify risks that are of practical interest. Although skill in the actual testing of such hypotheses is an acceptable by-product, the demonstrated skill set **must not** be confined to, or even largely focused on the actual construction of such tests. In particular, the student must

1. Be able to demonstrate a substantive understanding of “statistical significance,” and the sense in which p-values and null hypothesis testing offer a useful and practical articulation of risk assessment. To do this, the student must also be able to demonstrate mastery of the basic language of statistical experimental design and null hypothesis testing, and articulate the role that statistical modeling plays in the development and interpretation of “statistical significance.”
 2. Be able to articulate the strengths and weaknesses of using classical null hypothesis testing as a decision tool. Students should understand the sense in which common hypothesis testing, and the associated “significance” addressed in media, is intimately related to a perspective that looks for evidence against a claim, and infers about the truth of that claim based on the weight of that evidence
- C. (At least 20% of the course) - Evaluate common claims that arise from statistical constructs, like charts and graphs, tables and numerical summaries, through the important, but informal, act of human inference. Although skill in the actual construction of these constructs is an acceptable by-product, the demonstrated skill set **must not** be confined to, or even largely focused on these constructions. In particular, students must:
1. Be able to demonstrate an understanding of the challenges that confront informal inferences arising from these kinds of statistical entities and offer evidence that they can construct these inferences in a rational and informed manner.
 2. Be able to discuss the practical importance of effective conditional reasoning (e.g. false positives, Prosecutor’s paradox); the importance of hidden variables and confounding (e.g. Simpson’s paradox); the issue of association versus correlation and correlation and causation; the importance of having the right and/or enough information; and the problem of misinterpreting randomness.
- D. (At least 5% of the course) - Demonstrate information literacy by their measurable ability to independently identify and utilize appropriate information resources from a variety of sources. Instructors will collaborate with librarians to create a course-relevant component developing lifelong learning skills allowing students to identify, utilize, evaluate, apply and communicate information, a critical competency in becoming a contributing member of society.

The prerequisite for courses in this category is a course in the proposed category of “quantitative foundations.”

Guidelines for Course Designers

The ways in which the course outcomes are achieved, and the contexts in which the concepts are motivated, are the purview of individual departments, colleges, and instructors. However, while many of the concepts discussed in this course category are, at their root, complex mathematical concepts (e.g. the Central Limit Theorem), this course *is not* intended to be a mathematically complex course. Rather, the complexity of the course will likely be rooted in the ideas being discussed and the ways in which core concepts in statistical science connect to and surface in activities as common as reading the morning newspaper. With this in mind, the following suggestions are offered:

Curriculum-Embedded, Performance-Based Assessable Products

All students must create at least one assessable product that can be shared with the University's Assessment Office and the course syllabus must make clear what that product is. Individual instructors (or departments) are encouraged to consult with the Director of Assessment at the University, prior to the construction of a new syllabus. Rather than test knowledge or particular techniques, the assessment tool(s) should allow students to demonstrate an understanding of how statistical inference is used in decision making and to appraise the efficacy of statistical arguments that are reported for general consumption. That is, the assessment, too, should focus upon real world applications of learning outcomes A-D above. We recommend that the tool be validated, structured to allow electronic submission, and that an appropriate assessment rubric be developed based upon these criteria.