Class project 1:
Designing a Bayesian classifier

Objective

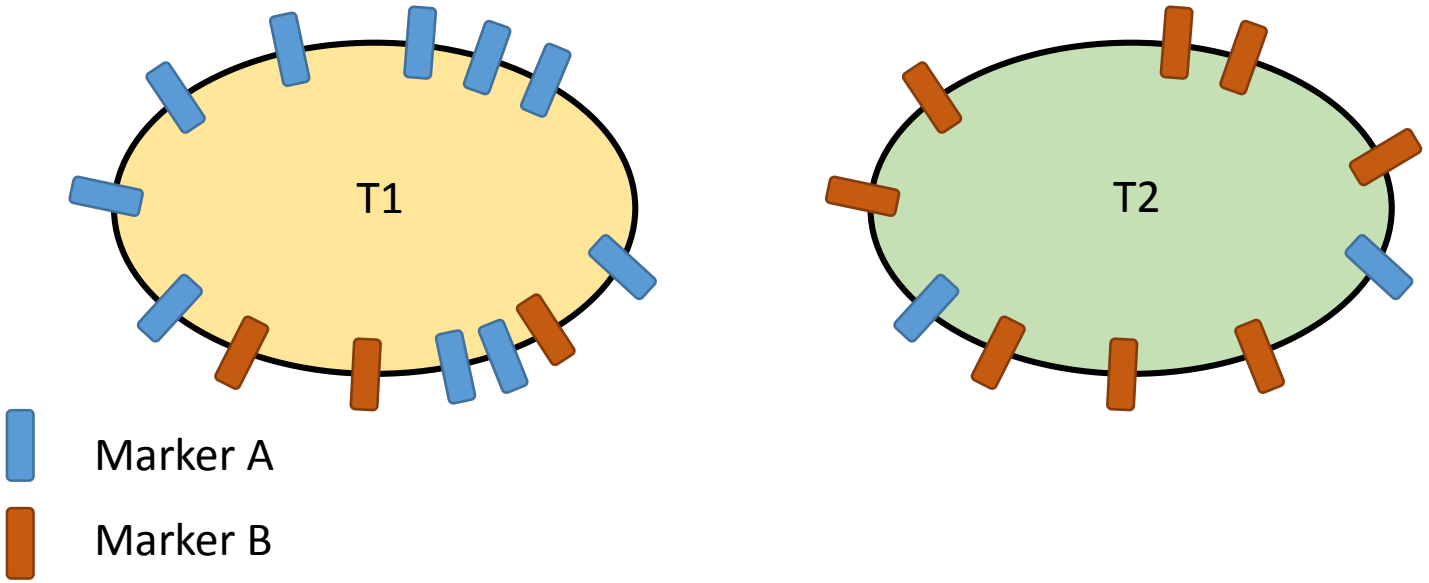
The objective of this project is to apply concepts of probability and statistics to drive and justify decisions in a design problem. Specifically, you will design a method for classifying entities (in this class T cells) in groups according to measurable properties subject to random variability.

Background

Cells express a large variety of surface molecules. The number of surface molecules varies from cell to cell in a random manner and their relative abundance is often cell-type dependent.

The complement of surface molecules is routinely used as markers to classify cells into major lineages (neurons, hepatocytes) or subtypes (effector T cells, memory T cells, etc.).

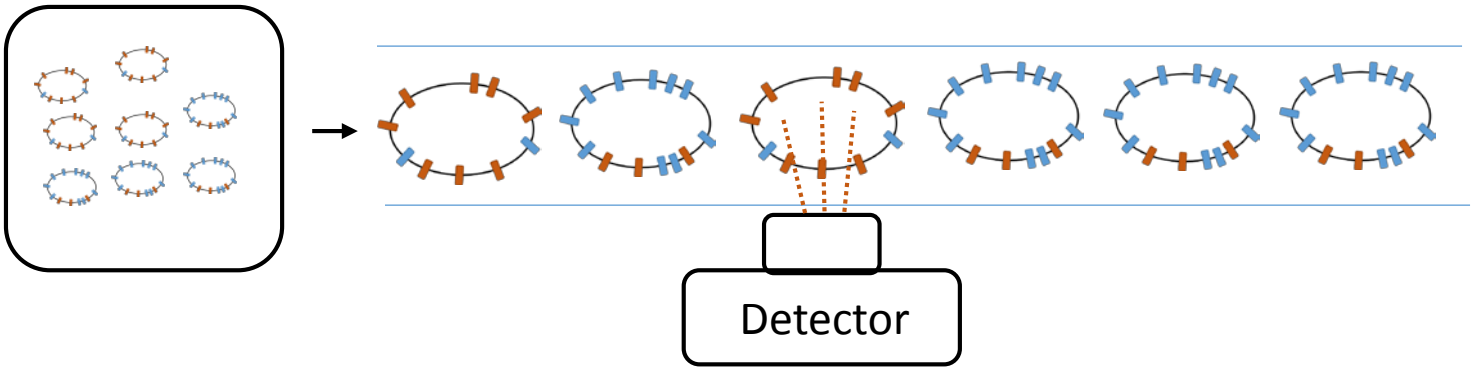
Many diagnostic tests rely on cell identification.



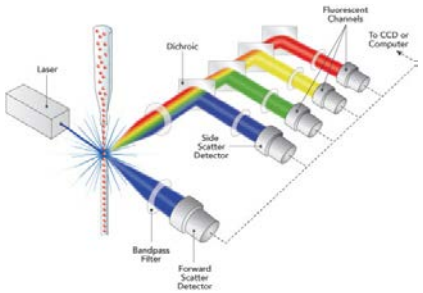
Background (cont.)

A flow cytometer is a device used to quickly analyze a large sample of cells treated with marker-specific fluorescent stains.

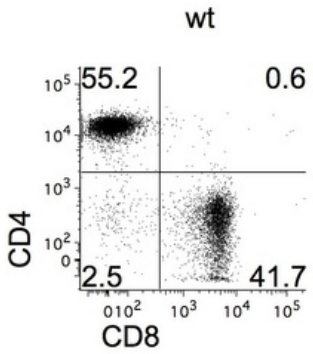
Cells move rapidly through a detector that quantifies fluorescence in various channels, each specific for a marker.



Beckman Coulter



https://www.semrock.com/flow-cytometry.aspx



Thiault et al, Nature Immunology 16, 628–638 (2015)

Goal

Designing a method for classifying cells based on the measured abundance of two markers (A and B) and their known joint PDFs in two cell types (T1 and T2).

To apply concepts of probability to drive and justify design choices as well as for evaluating the effectiveness of your method.

Deliverables

A written report describing your design process. The report must include quantitative support for each design choice. Figures should be included as necessary (or explicitly asked for). The report should address the effectiveness of your method as implemented by indicating the types and frequency (actual numbers) of classification errors.

Expectations

- Clearly written report that follows a logical path through the design process. All supporting calculations and requested items are present.
- The report clearly shows you understand concepts such as conditional and marginal probability, expectation, PDFs and CDFs and know how to apply them.
- Figures are numbered and labeled with title and caption. Figure are referenced in the text.
- All design decisions (even the ones we made for you) are explained and justified in terms of probability concepts with supporting calculations when adequate.
- Cogent discussion of improvements or alternative approaches.
- Reported is turned in on time.
- **Students abide by UT honor code. Evidence of plagiarism (e.g. presenting someone else's work) will be reported to SJS.**

Grading: 7.5 % of overall grade

1. *Organization and presentation of the report (0-5)*

   This critieria is about how you tell your design story.

   - 5- Well organized/presented.
   - 0- Poorly organized and/or incomplete.

2. *Understanding of concepts of probability (0-5)*

   This criteria is about mechanics; that is, doing things correctly.

   - 5- Reflects good understanding of the class material, in particular PDF's, conditional probability, marginal probability, expectation and variance, and Bayesian approaches.
   - 0- Incorrect calculations (e.g. using wrong expressions), results inconsistent with data or concepts from probability, failure to check for errors, and/or gross misinterpretation of intermediate results.

3. *Application of probability concepts to support design decisions (0 -5)*

   This criteria is about big picture and creativity.

   - 5- Demonstrates ability to transfer concepts from probability to the solution of a real-life problem.
   - 0- Student fail to frame the problem in terms of probability concepts or does it incorrectly.

Groups

- Students must form groups with peers registered in the same discussion session. Group size depends on which session you are attending.
  - Groups of 4 or 5 (sections Mon 3-4PM)
  - Groups of 3 or 4 (other sections)
- Contact me if you want to move out from a Mon 3-4 session.

- Ideally, same groups for project II.

- Students will have to present one project as a group at the end of the semester.

- Attendance to the group presentations (during discussion or lecture time) is mandatory. Presentations will be scheduled towards the end of the semester and indicated in advance.

Students must write their own reports. No group reports or "customized" group reports. Reports with similar wording will be considered academic dishonesty and referred to SJS.

We much prefer quality rather than quantity. Your report must reflect a rational design process. Avoid superfluous material.

We will be delighted to see creative solutions of your own. As long as you demonstrate mastery of the material and skills applying it, you will not be penalized for deviating from instructions in the hand-out.