Material Covered

Test 1 covers Sections 5.1–5.5

Format

The exam will include both an in-class portion and a take-home portion. The take-home portion will be due on Thursday, March 1.

Things to be sure you review

This is not intended to be an exhaustive list, but the items below are important things to review.

- Statistical models and distributions
  - You will be provided with the tables of distributions, so you don't need to memorize all that information, but you should know how to work with it.
  - You should recognize situtations when various favorite distribution families are (potentially) good models.
  - Basic probability rules and the cdf method
  - Recognizing a distribution from its kernel
  - Comparing data to a model (qq plots, density overlay on a histogram, etc.)
  - Mixture distributions (as we used in the Old Faithful example)
- Point Estimates for Parameters in a Model
  - Method of moments estimation
  - Definition of likelihood function
  - Finding MLEs by maximizing the likelihood or log-likelihood function (analytically and/or numerically)
  - For numerical results, you should know how to (a) create log-likelihood functions in R and how to use the `maxLik` package to optimize.
- Likelihood Ratio Tests
  - Definition and calculation of the likelihood ratio statistic (λ) and the normalized statistic $W = -2\log(\lambda)$.
  - (Approximate) distribution of likeihood ratio statistic
    - using $\dim \Omega$ and $\dim \Omega_0$ to calculate degrees of freedom
    - when the Chi-squared approximation is better or worse
    - using simulations to approximation the null distribution of the test statistic.
- Confidence Intervals
  - Score and information functions
  - Quadratic approximation to the log-liklihood function
  - Wald and likelihood ratio confidence intervals
    - computing (approximate) standard error using the observed information.
    - Inverting p-values to obtain confidence intervals

R skills

You should be able to do the following in R:

- Get data into R.
- Prepare documents using RMarkdown.
- Do probability calcuations for the distributions we have seen (or any others that work similarly) using functions like `dbinom()`, `pbinom()`, `qbinom()`, and `rbinom()`.
- Define functions and make sure they are vectorized as needed for you application.
- Use `maxLik()` to maximize log-likelihood functions.
- Use `uniroot()` to find roots of a function. (Example application: likelihood confidence intervals.)
- Use `do()` or `expand.grid()` to perform simulations. `expand.grid()` can also be used for grid search.
- Make various plots (typically using `ggformula`) including histograms, scatter plots, qq-plots, plots of pdfs (`gf_dist()`), and plots of your own functions.

Further suggestions

- No mystery numbers allowed. It should be clear where every number comes from.

  When you create objects in R by a sequence of operations, show me the intermediate values as you go along. The following coding paradigm is useful:

```
# this both stores your intermediate value and displays it
x <- 3 + 5; x

## [1] 8
```

- Use notation well.
  - You are required to understand and use the notation we have introduced in class. This includes correct use of the equals sign (=).
  - Introduce the notation you use and what it stands for. (Example: "Let $\boldsymbol{X} \overset{\text{iid}}{\sim} \mathsf{Norm}(\mu, \sigma)$", and let $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$. Then $Z \sim \mathsf{Norm}(0, 1)$.)
  - Remember that we sometimes use shorthand notation to focus our attention on the most important things. We might write $L(\theta)$ instead of $L(\theta; \boldsymbol{x})$ or $\sum \theta \log(x)$ instead of $\sum_{i=1}^{n} \theta \log(x_i)$. You should always know what the fully verbose version would be and what each letter and symbol in the notation stands for. When in doubt, be more verbose rather than less verbose with your notation.
  - You may invent notation as long as you explain it.

- Don't be afraid to use words.

  In any case, do your work in "paragraph order" (left to right, top to bottom).