Probability/Preparation

- basic notions
  - {not, and, or}; conditional probability; independence
  - random variables: X, $f(x)$, $F(x)$, $\mathbb{E}(X)$, $\mathbb{V}\text{ar}(X)$
- harder ideas (takes time and experience to master)
  - $f_{XY}(x, y) \to f_X(x) \to f_{Y|X}(y|x)$; covariance
  - law of total $\{\mathbb{P}(...), \mathbb{E}(...), \mathbb{V}\text{ar}(...)\}$
- two specific problems (important for statistics)
  - distributions of $g(X, Y)$, $h(X_1, X_2, ..., X_n)$
  - compound distributions (hierarchically specified ... ... models)

Frequentist Estimation

- estimator
  - a random variable in itself
  - referring to the procedure of estimation
  - rather than the numeric value
- finding estimators
  - method of moments; maximum likelihood
- assessing estimators
  - sampling distribution
  - bias, variance, mean squared error

Bayesian Estimation

- basic set-up:
  - $\pi(\theta|\text{data}) = f(\text{data}|\theta)\pi(\theta)/m(\text{data})$
- simultaneous estimation problem: (h for "hyperparameter")
  - $X_i \sim f(\cdot|\theta_i)$ independently for $i = 1, ..., n$, and $\theta_i \overset{iid}{\sim} \pi(\cdot|h)$
- simple Bayes: (h non-random, specified)
  - specify h, and use $\pi(\theta_i|x_i, h) = f(x_i|\theta_i)\pi(\theta_i|h)/m(x_i|h)$
- empirical Bayes: (h non-random, estimated)
  - estimate h w/ all $X_i \overset{iid}{\sim} m(\cdot|h)$, and plug into $\pi(\theta_i|x_i, h)$
- hierarchical Bayes: (h random)
  - put prior on h, and go for $\pi(\theta_1, ..., \theta_n, h|x_1, ..., x_n)$

Confidence Sets

- pivotal quantity $h(\theta; \boldsymbol{D})$
  - must know its distribution (not easy)
  - e.g., $T(\mu; \boldsymbol{D}) \equiv (\bar{X} - \mu)/\sqrt{S^2/n}$
- e.g., likelihood ratio $\Lambda(\theta; \boldsymbol{D})$
  - $2 \log \Lambda(\theta; \boldsymbol{D})$, and the χ^2-distribution

Significance Tests

- key jargon
  - null hypothesis, test statistic, significance level, p-value
- test statistic vs. pivotal quantity
  - $\underbrace{\mathbb{P}[h(\theta; \boldsymbol{D}) \in C_\alpha] = 1 - \alpha}_{\{\theta:\ h(\theta;\boldsymbol{D}) \in C_\alpha\} \text{ is a confidence set}} \Rightarrow \underbrace{\mathbb{P}[h(\theta_0; \boldsymbol{D}) \notin C_\alpha]}_{\theta_0 \text{ rejected } \notin \text{confidence set}} \stackrel{H_0}{=} \alpha$
- some specific tests
  - likelihood-ratio test (LRT), t-test, (later) F-test
- multiple testing problems
  - Bonferroni, Benjamini-Hockberg

Computational Techniques

- EM algorithm (basically, a "fancy" MLE algorithm)
  - useful for problems with missing/latent data
- Gibbs sampler
  - sequentially draw from conditional distributions
  - useful for intractable posterior distributions
- Bootstrap
  - $h(\widehat{\theta}; \boldsymbol{D}^{*(1)}),\ h(\widehat{\theta}; \boldsymbol{D}^{*(2)}),\ \ldots,\ h(\widehat{\theta}; \boldsymbol{D}^{*(B)})$
  - approximates the distribution of $h(\theta; \boldsymbol{D})$
  - versatile, can replicate "anything" we do on $\boldsymbol{D}$

Linear Models

- in terms of ideas, nothing new
  - e.g., MLE, bias, variance, testing (t, F)
- in terms of techniques (i.e., math), mostly linear algebra
  - span($\boldsymbol{X}$) $\subset \mathbb{R}^n$, projection, hat matrix $\boldsymbol{H}$
- model assessment and choice
  - F-test (nested models), CV, GCV, AIC
- ridge regression
  - variance reduction, multicollinearity
- generalized linear models (GLMs)
  - exponential families, Newton-Raphson (IRLS)

Experimental Design

- some important ideas + jargon
  - observational vs. experimental studies
  - association vs. causation
  - confounding, randomization, blocking
- in terms of techniques, nothing new (again)
  - ANOVA; F-tests — just nested linear models

Survey Sampling

- in terms of ideas, nothing new (again)
  - e.g., bias, variance
- in terms of techniques, just one new-ish "trick"
  - $I(x_j \in \mathcal{S})$, $I(x_j, x_\ell \in \mathcal{S})$; $\mathbb{P}(x_j \in \mathcal{S})$, $\mathbb{P}(x_j, x_\ell \in \mathcal{S})$
  - Horvitz-Thompson, inverse probability weighting
- in terms of methods, three basic strategies
  - simple random sampling
  - stratified sampling
  - cluster sampling (brief mention only)

Final Examination

(December 8, 2023; 16:00 – 18:30)

- Q1 (25 pts): matching and multiple choices on various topics
- Q2 (15 pts): Bayesian inference
- Q3 (18 pts): EM algorithm
- Q4 (20 pts): linear model + significance tests
- Q5 (12 pts): confidence interval
- Q6 (10 pts): survey sampling

Remark

Questions are NOT ordered by difficulty—e.g., for Q1 the questions are arranged by spacing and layout considerations alone.

Office Hours Next Week

Monday, December 4, 2023

12:00 – 13:00; M3 2103B
(usual)

14:00 – 18:00; M3 4007†
(extra)

Wednesday, December 6, 2023

10:00 – 16:00; M3 4007†
(extra)

†with some coffee breaks in between

Feedback

- official UW course feedback
  - completely anonymous
  - `https://perceptions.uwaterloo.ca`
  - by 23:59 EST, Tuesday, December 5, 2023
- *Essential Statistics* book
  - please consider writing a review
  - comments, errors, typos, ... please email