Lecture 23

Some remarks on Bayesian statistics
All the statistics we've done so far in this course are sometimes referred to as "frequentist" statistics. The basic idea is that we collect data, and then we make an inference on that data without making prior assumptions about the conclusion. It has been the most popular form of statistics for quite some time, but there are critiques of the process. There is a lot of discussion about P-values, for instance, and the consequences of publishing only affirmative results that can be misleading. It's worth considering these controversies when analyzing studies.

But there is another approach to statistics that comes at some of these issues from a completely different perspective. Bayesian statistics is founded in the idea of conditional probability and Bayes' Rule (which we've covered in this class). Recall:

$$P(A \mid B) = \frac{P(A \,\&\, B)}{P(B)}.$$

Bayesian statistics is a branch of statistics that deals with the analysis of uncertain events or phenomena using probability theory. It is named after Thomas Bayes, an 18th-century British mathematician, who introduced the concept of conditional probability and developed the foundations of Bayesian inference.

In Bayesian statistics, prior information is taken into account as we take in new data. Typically, some prior distribution is assumed, and the new data is taken in to modify the distribution, and this process continues each time new data is assimilated. This can be useful if you have reasons to believe something about the result (say for physical reasons, from differential equations or other areas of mathematics and science), then that information can be incorporated into the analysis. New data can modify these prior assumptions, and lots of data can modify them a lot. Many Bayesians argue that this process is more similar to the way that people actually reason. The reference [4] linked below is an entire online text about Bayesian statistics and they have a more in-depth discussion of the difference between Bayesian and frequentist statistics.

In our expression above, $P(A)$ and $P(B)$ are the prior probabilities of A and B, respectively, and $P(A|B)$ is the likelihood of A given B.

Bayesian statistics provides a framework for updating beliefs based on new evidence. It allows for the incorporation of prior knowledge, which can be subjective or based on previous data or studies. As new data becomes available, the prior beliefs are updated to form the posterior beliefs, which provide a more accurate representation of the uncertainty.

It is worth noting, however, that while the philosophy of Bayesian statistics and frequentist approaches are quite different, and lead to different computational methods, nonetheless, in the presence of sufficient data, the results of the analysis converge to the same numerical solutions.

One of the advantages of Bayesian statistics is its flexibility in handling complex problems. It allows for the incorporation of various sources of uncertainty and can handle small sample sizes effectively.

Bayesian methods are particularly useful when dealing with parameter estimation, hypothesis testing, decision-making under uncertainty, and predictive modeling.

Bayesian statistics has applications in various fields, including medicine, finance, engineering, and machine learning. It is used in Bayesian networks, Bayesian hierarchical models, Bayesian regression, and Bayesian decision theory, among others. Markov Chain Monte Carlo (MCMC) methods, such as Gibbs sampling and Metropolis-Hastings algorithm, are often employed to approximate the posterior distribution when analytical solutions are not feasible.

Overall, Bayesian statistics offers a coherent framework for reasoning under uncertainty, allowing for the integration of prior knowledge and observed data to make informed decisions and draw robust conclusions.

Bayesian statistics can be applied in various ways across different fields. Here are a couple of examples illustrating its practical applications:

Drug Efficacy Testing: Suppose a pharmaceutical company wants to test the effectiveness of a new drug. They conduct a clinical trial with a sample of patients and collect data on whether the drug successfully treats the condition. Bayesian statistics can be used to analyze the results and make inferences. Prior to the trial, the company might have some initial beliefs about the drug's efficacy based on previous studies or expert opinions, which can be represented as a prior probability distribution. By combining this prior distribution with the observed data from the trial, Bayesian analysis can yield a posterior distribution that provides updated information on the drug's effectiveness. This posterior distribution can then be used to estimate the probability that the drug is effective and make decisions about its future development.

Fraud Detection: In the field of fraud detection, Bayesian statistics can be applied to identify potentially fraudulent activities. For example, in credit card fraud detection, a Bayesian approach can be used to calculate the probability that a given transaction is fraudulent based on various factors such as transaction amount, location, time, and past transaction history. Prior probabilities can be derived from historical data on fraudulent and non-fraudulent transactions. By combining these priors with the observed characteristics of a new transaction, the Bayesian framework can update the probabilities and provide a posterior probability of fraud. If the posterior probability exceeds a certain threshold, further investigation or action can be taken.

These examples demonstrate how Bayesian statistics enables the incorporation of prior knowledge, subjective beliefs, and observed data to make informed decisions and draw reliable conclusions. By updating probabilities through Bayesian inference, it allows for continuous learning and updating of beliefs as new evidence becomes available.

Bayesian statistics can be applied to probability distributions in various ways, allowing for the estimation, updating, and inference of parameters within these distributions. Here are a few ways Bayesian methods are applied to probability distributions:

Parameter Estimation: In frequentist statistics, point estimates of parameters are often obtained using methods such as maximum likelihood estimation (MLE). In Bayesian statistics, however, parameter estimation involves obtaining the posterior distribution of the parameters given the data. This is accomplished by specifying a prior distribution that represents our beliefs about the parameters before observing any data. The prior distribution is then combined with the likelihood function, which represents the probability of the observed data given the parameters, using Bayes' theorem. The resulting posterior distribution provides updated information about the parameters, incorporating both the prior beliefs and the observed data.

Model Selection and Comparison: Bayesian methods also facilitate model selection and comparison by considering probability distributions over models. This is particularly useful when comparing complex models with different numbers of parameters or different structural assumptions. Bayesian model selection involves calculating the posterior probabilities of different models given the observed data. The models' prior probabilities, often based on their complexity or prior knowledge, are combined with the likelihoods of the data given each model to obtain the posterior probabilities. The model with the highest posterior probability is then considered the most plausible given the data.

Prediction and Forecasting: Bayesian statistics enables probabilistic predictions and forecasting by incorporating uncertainty through probability distributions. Rather than providing a single point estimate, Bayesian methods yield a predictive distribution that quantifies the uncertainty associated with the prediction. This is achieved by combining prior beliefs about the parameters or future data with the observed data using Bayes' theorem. The predictive distribution represents the updated knowledge or belief about the future outcomes, accounting for both the prior information and the observed data.

In all these applications, Bayesian statistics allows for the representation of uncertainty through probability distributions. It provides a coherent framework for reasoning about and updating knowledge based on prior beliefs and observed data, resulting in posterior distributions that incorporate both sources of information. By using probability distributions, Bayesian methods provide a more comprehensive and flexible approach to statistical inference and analysis.

In the context of hypothesis testing, a Bayesian approach provides an alternative to the classical frequentist approach. While the frequentist approach focuses on the long-run behavior of statistical procedures, the Bayesian approach directly incorporates prior beliefs and updates them based on observed data to obtain posterior probabilities. Here's how a Bayesian approach to hypothesis testing typically works:

Specify the Hypotheses: In Bayesian hypothesis testing, you start by specifying the competing hypotheses. The hypotheses can be represented as probability distributions over the parameters of interest. For example, you might have a null hypothesis ($H0$) and an alternative hypothesis ($H1$), each defined by a probability distribution that captures uncertainty about the parameters under consideration.

Assign Prior Probabilities: Next, you need to assign prior probabilities to the competing hypotheses. These prior probabilities represent your initial beliefs or subjective judgments about the likelihood of each hypothesis being true. The priors can be based on previous data, expert opinions, or non-informative priors that spread the probability mass evenly.

Collect and Analyze Data: After specifying the hypotheses and priors, you collect and analyze the data. The observed data are used to update the prior probabilities, resulting in posterior probabilities that reflect your updated beliefs about the hypotheses.

Compute the Bayes Factor or Posterior Odds: To compare the strength of evidence between the hypotheses, you can compute the Bayes factor or posterior odds. The Bayes factor quantifies the relative support for one hypothesis compared to another by evaluating the ratio of the likelihoods under the competing hypotheses. Alternatively, the posterior odds compares the probabilities of the hypotheses based on the posterior distributions.

Make a Decision: Based on the computed Bayes factor or posterior odds, you can make a decision about which hypothesis is more plausible. A common decision rule is to select the hypothesis with the highest posterior probability or the hypothesis with a predefined threshold for substantial evidence. In some cases, a decision can also be made by considering the expected utility or loss associated with different decisions.

The Bayesian approach to hypothesis testing allows for the incorporation of prior beliefs, which can be particularly useful in situations with limited data. It provides a more intuitive interpretation of the evidence and allows for continuous updating of beliefs as new data becomes available. However, it's important to note that prior specifications can heavily influence the results, and subjective choices in assigning priors need to be carefully justified and transparently reported.

Bayesian statistics and frequentist approaches have distinct advantages and disadvantages. Here's a comparison of the two:

Advantages of Bayesian Statistics:
Incorporation of Prior Knowledge: Bayesian statistics allows for the explicit incorporation of prior beliefs or knowledge about the parameters or hypotheses of interest. This is particularly useful when dealing with limited data or when expert opinions are available.

Uncertainty Quantification: Bayesian methods provide a natural way to quantify uncertainty through probability distributions. The posterior distribution represents a comprehensive summary of the uncertainty in the parameters or predictions, accounting for both the prior information and the observed data.

Flexibility with Small Sample Sizes: Bayesian approaches can yield more stable and reliable estimates with small sample sizes by incorporating prior information. The prior distribution acts as a regularizer, providing a smoothing effect and reducing the impact of outliers or extreme observations.

Coherent Decision-Making Framework: Bayesian statistics offers a coherent decision-making framework by directly providing posterior probabilities. These probabilities can be used to compare hypotheses, make decisions, and perform cost-benefit analyses based on the expected utility or loss.

Disadvantages of Bayesian Statistics:
Subjectivity in Prior Specification: The choice of prior distributions in Bayesian analysis introduces subjectivity, as different analysts may have different prior beliefs. This can raise concerns about the objectivity and reproducibility of the results. Sensitivity analyses and robustness checks are important to assess the impact of prior choices.

Computational Complexity: Bayesian methods often involve complex calculations, especially when dealing with high-dimensional models or complex likelihood functions. Approximation techniques like MCMC methods can be computationally intensive and time-consuming.

Interpretation Challenges: Bayesian results are often expressed in terms of probability distributions, which may be more challenging to interpret for some stakeholders compared to point estimates or p-values used in frequentist approaches. Clear communication of results is crucial to avoid misinterpretation.

Advantages of Frequentist Approaches:

Strong Foundation in Statistical Theory: Frequentist approaches have a well-established theoretical foundation and are often associated with rigorous statistical inference. They provide asymptotic guarantees and maintain certain properties, such as unbiasedness or consistency, under certain assumptions.

Objective Decision Rules: Frequentist approaches provide clear decision rules, such as rejecting or not rejecting a null hypothesis based on pre-defined significance levels. This can simplify decision-making and facilitate comparisons across different studies or researchers.

Disadvantages of Frequentist Approaches:

Limited Treatment of Uncertainty: Frequentist methods typically provide point estimates or confidence intervals that only capture sampling variability, without directly quantifying other sources of uncertainty or incorporating prior knowledge.

Lack of Flexibility with Small Sample Sizes: Frequentist methods may yield unreliable estimates or results with small sample sizes, as they heavily rely on the observed data without the regularization effect of prior information.

Difficulty Handling Complex Models: Frequentist approaches can be less flexible when dealing with complex models or when incorporating complex prior information. They may require large sample sizes or rely on asymptotic approximations that may not hold in small samples.

In practice, the choice between Bayesian and frequentist approaches depends on various factors, including the availability of prior information, the specific research question, the nature of the data, and the preferences of the researcher or analyst. Both approaches have their strengths and weaknesses, and the choice should be based on a careful consideration of the specific context and requirements of the analysis.

Next semester, we'll continue with the frequentist approach, but we may touch on a couple of techniques that were born from Bayesian approaches.

Review for final exam

Review material from Exam #1 and Exam #2 for the comprehensive portion of the Final.

For the recent material, focus on the following:

- Non-parametric statistics
    - Wilcoxon test for sign test/rank-sum test
    - Non-parametric ANOVA (Kruskal-Wallis, Friedman's ANOVA)

    - Permutation tests
    - Bootstrapping
- χ^2 tests
    - Goodness-of-fit tests
    - Test of homogeneity
    - Test of independence
    - Fisher Exact Test

The final exam will have the same format as the two previous exams.

Don't forget to submit the final draft of your final project as well. Be sure to include suggested changes from your rough draft into your final version.

References:

1. https://assets.openstax.org/oscms-prodcms/media/documents/IntroductoryStatistics-OP_i6tAI7e.pdf
2. https://faculty.ksu.edu.sa/sites/default/files/probability_and_statistics_for_engineering_and_the_sciences.pdf
3. http://www.scholarpedia.org/article/Bayesian_statistics
4. https://statswithr.github.io/book/the-basics-of-bayesian-statistics.html