Comparison of Bayesian and Frequentist Inference

18.05 Spring 2022

Probability (mathematics)

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

Everyone uses Bayes' formula when the prior $P(H)$ is known.

Bayesian path

Frequentist path

Statistics (art)

$$P_{\text{Posterior}}(H|D) = \frac{P(D|H)P_{\text{prior}}(H)}{P(D)}$$

Bayesians require a prior, so they develop one from the best information they have.

Likelihood $L(H; D) = P(D|H)$

Without a known prior, frequentists draw inferences from just the likelihood function.

Announcements/Agenda

Announcements

- Sarah has OH after class for people with pset questions.
- Exam 2 is a week from Thursday.
- Practice materials will be up this week.
- Next Tuesday, class will be for review.

Agenda

- Comparison of Bayesian and Frequentist inference
- Go over some of last week's board problems.

Compare Bayesian and Frequentist inference

Bayesian inference

- Uses priors
- Logically impeccable
- Probabilities can be interpreted
- Prior is subjective

Frequentist inference

- No prior
- Objective –everyone gets the same answer
- Logically complex
- Conditional probability of error is often misinterpreted as total probability of error
- Requires complete description of experimental protocol and data analysis protocol before starting the experiment. (This is both good and bad)

Concept question: Significance tests

Three different tests are run, all with significance level $\alpha = 0.05$.

Experiment 1: finds $p = 0.003$ and rejects its null hypothesis H_0.

Experiment 2: finds $p = 0.049$ and rejects its null hypothesis.

Experiment 3: finds $p = 0.15$ and fails to rejects its null hypothesis.

Which result has the highest probability of being correct?

1. Experiment 1
2. Experiment 2
3. Experiment 3
4. Impossible to say.

Board question: Stop!

Experiments are run to test a coin that is suspected of being biased towards heads. The significance level is set to $\alpha = 0.1$
Experiment 1: Toss a coin 5 times. Report the sequence of tosses.

Experiment 2: Toss a coin until the first tails. Report the sequence of tosses.

(a) Give the test statistic, null distribution and rejection region for each experiment. List all sequences of tosses that produce a test statistic in the rejection region for each experiment.

(b) Suppose the data is $HHHHT$.

(i) Do the significance test for both types of experiment.

(ii) Do a Bayesian update starting from a flat prior: Beta(1,1).

Draw some conclusions about the fairness of coin from your posterior.
(Use R: `pbeta` for computation in part (b).)

Board question: Stop II

For each of the following experiments (all done with $\alpha = 0.05$)
(a) Comment on the validity of the claims.
(b) Find the true probability of a type I error in each experimental setup.

1. Experiment 1. By design Alessandre did 50 trials and computed $p = 0.04$.
   They report $p = 0.04$ with $n = 50$ and declare it significant.
2. Experiment 2. Sara did 50 trials and computed $p = 0.06$.
   Since this was not significant, she then did 50 more trials and computed $p = 0.04$ based on all 100 trials.
   She reports $p = 0.04$ with $n = 100$ and declares it significant.
3. Experiment 3. Gabriel did 50 trials and computed $p = 0.06$.
   Since this was not significant, he started over and computed $p = 0.04$ based on the next 50 trials.
   He reports $p = 0.04$ with $n = 50$ and declares it statistically significant.

From Class 18: Type I errors Q1

Suppose a journal will only publish results that are statistically significant at the 0.05 level. What percentage of the papers it publishes contain type I errors?

From Class 18: Type I errors Q2

Jerry desperately wants to cure diseases but he is terrible at designing effective treatments. He is however a careful scientist and statistician, so he randomly divides his patients into control and treatment groups. The control group gets a placebo and the treatment group gets the experimental treatment. His null hypothesis H_0 is that the treatment is no better than the placebo. He uses a significance level of $\alpha = 0.05$. If his p-value is less than α he publishes a paper claiming the treatment is significantly better than a placebo.

(a) Since his treatments are never, in fact, effective what percentage of his experiments result in published papers?

(b) What percentage of his published papers contain type I errors, i.e. describe treatments that are no better than placebo?

From Class 18: Type I errors: Q3

Jen is a genius at designing treatments, so all of her proposed treatments are effective. She is also a careful scientist and statistician, so she too runs double-blind, placebo controlled, randomized studies. Her null hypothesis is always that the new treatment is no better than the placebo. She also uses a significance level of $\alpha = 0.05$ and publishes a paper if $p < \alpha$.

(a) How could you determine what percentage of her experiments result in publications?

(b) What percentage of her published papers contain type I errors, i.e. describe treatments that are, in fact, no better than placebo?

From Class 19: chi-square for independence

(From Rice, *Mathematical Statistics and Data Analysis*, 2nd ed. p.489)

Consider the following contingency table of counts

Education	Married once	Married multiple times	Total
College	550	61	611
No college	681	144	825
Total	1231	205	1436

Use a chi-square test with significance level 0.01 to test the hypothesis that the number of marriages and education level are independent.