STAT 101
Dr. Kari Lock Morgan

Bayesian Inference

SECTION 11.1, 11.2

- Bayes rule (11.2)
- Bayesian inference (not in book)

Review

What is the definition of the p-value?

a) P(statistic as extreme as that observed if H_0 is true)
b) P(H_0 is true if statistic as extreme as that observed)

Breast Cancer Screening

A 40-year old woman participates in routine screening and has a positive mammography. What's the probability she has cancer?

a) 0-10%
b) 10-25%
c) 25-50%
d) 50-75%
e) 75-100%

Breast Cancer Screening

- 1% of women at age 40 who participate in routine screening have breast cancer.
- 80% of women with breast cancer get positive mammographies.
- 9.6% of women without breast cancer get positive mammographies.
- A 40-year old woman participates in routine screening and has a positive mammography. What's the probability she has cancer?

Breast Cancer Screening

A 40-year old woman participates in routine screening and has a positive mammography. What's the probability she has cancer?

a) 0-10%
b) 10-25%
c) 25-50%
d) 50-75%
e) 75-100%

Breast Cancer Screening

A 40-year old woman participates in routine screening and has a positive mammography. What's the probability she has cancer?

What is this asking for?

a) P(cancer if positive mammography)
b) P(positive mammography if cancer)
c) P(positive mammography if no cancer)
d) P(positive mammography)
e) P(cancer)

Bayes Rule

- We know P(positive mammography if cancer)… how do we get to P(cancer if positive mammography)?
- How do we go from P(A if B) to P(B if A)?

$$P(A \text{ if } B) = \frac{P(B \text{ if } A)P(A)}{P(B)}$$

Bayes Rule

$$P(A \text{ if } B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$\Rightarrow P(A \text{ and } B) = P(A \text{ if } B)P(B)$$

$$P(A \text{ and } B) = P(B \text{ if } A)P(A)$$

$$\Rightarrow P(A \text{ if } B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{P(B \text{ if } A)P(A)}{P(B)}$$

$$\Rightarrow P(A \text{ if } B) = \frac{P(B \text{ if } A)P(A)}{P(B)}$$ <- **Bayes Rule**

Rev. Thomas Bayes

1702 - 1761

Breast Cancer Screening

$$P(\text{cancer if positive}) = \frac{P(\text{positive if cancer})P(\text{cancer})}{P(\text{positive})}$$

- 1% of women at age 40 who participate in routine screening have breast cancer.
- 80% of women with breast cancer get positive mammographies.
- 9.6% of women without breast cancer get positive mammographies.

P(positive)

- How do we figure out P(positive)?
- We know:
  - 80% of women with breast cancer get positive mammographies.
  - 9.6% of women without breast cancer get positive mammographies.
- We need to average these two numbers, weighted by the proportion of people with breast cancer:

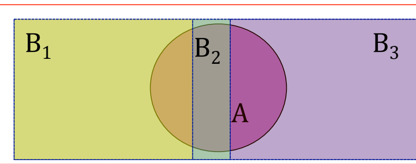
$$P(\text{positive if cancer})P(\text{cancer}) + P(\text{positive if no cancer})P(\text{no cancer})$$

$$0.8 \times P(\text{cancer}) + 0.096 \times P(\text{no cancer})$$

Law of Total Probability

- If events B_1 through B_k are *disjoint* and together make up all possibilities, then

$$P(A) = P(A \text{ and } B_1) + P(A \text{ and } B_2) + \ldots + P(A \text{ and } B_k)$$

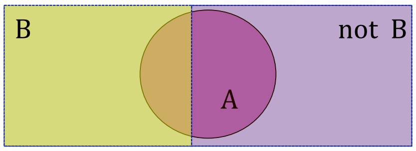


$$P(A) = P(A \text{ if } B_1)P(B_1) + P(A \text{ if } B_2)P(B_2) + \ldots + P(A \text{ if } B_k)P(B_k)$$

Law of Total Probability

- Special case: B and not B

$$P(A) = P(A \text{ and } B) + P(A \text{ and not B})$$



$$P(A) = P(A \text{ if } B)P(B) + P(A \text{ if not } B)P(\text{not } B)$$

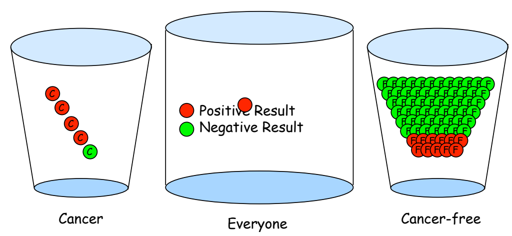
P(positive)

Use the law of total probability to find P(positive).

Breast Cancer Screening

$$P(\text{cancer if positive}) = \frac{P(\text{positive if cancer})P(\text{cancer})}{P(\text{positive})}$$

- 1% of women at age 40 who participate in routine screening have breast cancer.
- 80% of women with breast cancer get positive mammographies.
- 9.6% of women without breast cancer get positive mammographies.



- We randomly pick a ball from the Everyone bin.
- If the ball is red/positive, is it more likely to be from the Cancer or Cancer-free bin?

100,000 women in the population

- 1% → 1000 have cancer
  - 80% → 800 test positive
  - 20% → 200 test negative
- 99% → 99,000 cancer-free
  - 9.6% → 9,504 test positive
  - 90.4% → 89,496 test negative

Thus, 800/(800+9,504) = 7.8% of positive results have cancer

Hypotheses

H_0 : no cancer
H_a : cancer

Data: positive mammography

p-value = P(statistic as extreme as observed if H_0 true)
= P(positive mammography if no cancer)
= 0.096

The probability of getting a positive mammography just by random chance, if the woman does not have cancer, is 0.096.

Hypotheses

H_0 : no cancer
H_a : cancer

Data: positive mammography

You don't really want the p-value, you want the probability that the woman has cancer!

You want P(H_0 true if data), not P(data if H_0 true)

Hypotheses

H_0 : no cancer
H_a : cancer

Data: positive mammography

Using Bayes Rule:
P(H_a true if data) = P(cancer if data) = 0.078
P(H_0 true if data) = P(no cancer | data) = 0.922

This tells a very different story than a p-value of 0.096!

Frequentist Inference

- ***Frequentist Inference*** considers what would happen if the data collection process (sampling or experiment) was repeated many times
- Probability is considered to be the proportion of times an event would happen if repeated many times
- In frequentist inference, we condition on some unknown truth, and find the probability of our data given this unknown truth

Frequentist Inference

- Everything we have done so far in class is based on frequentist inference
- A confidence interval is created to capture the truth for a specified proportion of all samples
- A p-value is the proportion of times you would get results as extreme as those observed, if the null hypothesis were true

Bayesian Inference

- ***Bayesian inference*** does not think about repeated sampling or repeating the experiment, but only what you can tell from your single observed data set
- Probability is considered to be the subjective degree of belief in some statement
- In Bayesian inference we condition on the data, and find the probability of some unknown parameter, given the data

Fixed and Random

- In frequentist inference, the parameter is considered fixed and the sample statistic is random
- In Bayesian inference, the statistic is considered fixed, and the parameter is considered random

Bayesian Inference

Frequentist: P(data if truth)

Bayesian: P(truth if data)

- How are they connected?

$$P(truth \text{ if } data) = \frac{P(data \text{ if } truth)P(truth)}{P(data)}$$

Bayesian Inference

POSTERIOR Probability

PRIOR Probability

$$P(truth \text{ if } data) = \frac{P(data \text{ if } truth)P(truth)}{P(data)}$$

- Prior probability: probability of a statement being true, before looking at the data
- Posterior probability: probability of the statement being true, after updating the prior probability based on the data

Breast Cancer

- Before getting the positive result from her mammography, the prior probability that the woman has breast cancer is 1%
- Given data (the positive mammography), update this probability using Bayes rule:

$$\frac{P(data \text{ if } truth)P(truth)}{P(data)} = \frac{0.8 \times 0.01}{0.103} = 0.078$$

- The posterior probability of her having breast cancer is 0.078.

Paternity

- A woman is pregnant. However, she slept with two different guys (call them Al and Bob) close to the time of conception, and does not know who the father is.
- What is the prior probability that Al is the father?
- The baby is born with blue eyes. Al has brown eyes and Bob has blue eyes. Update based on this information to find the posterior probability that Al is the father.

Eye Color

- In reality eye color comes from several genes, and there are several possibilities but let's simplify here:
  - Brown is dominant, blue is recessive
  - One gene comes from each parent
  - BB, bB, Bb would all result in brown eyes
  - Only bb results in blue eyes
- To make it a bit easier: You know that Al's mother and the mother of the child both have blue eyes.

Paternity

What is the probability that Al is the father?

a) 1/2
b) 1/3
c) 1/4
d) 1/5
e) No idea

Paternity

Bayesian Inference

- Why isn't everyone a Bayesian? ???

$$P(truth \text{ if } data) = \frac{P(data \text{ if } truth)P(truth)}{P(data)}$$

- Need some "prior belief" for the probability of the truth
- Also, until recently, it was hard to be a Bayesian (needed complicated math.) Now, we can let computers do the work for us!

Inference

Both kinds of inference have the same goal, and it is a goal fundamental to statistics:

to use information from the data to gain information about the unknown truth

To Do

- Read 11.2
- Do Project 2 (due Wednesday, 4/23)
- Do Homework 9 (due Wednesday, 4/23)

Course Evaluations