Lecture 29: Bayesian Statistics

Chapter 2.2.7

Recall Conditional Probability

New Notation

Previously

Previously

So recall from previously we have the following 2×2 table of possible outcomes:

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	$(1-0.05) \times 900 = 855$	$0.05 \times 900 = 45$
	H_A true	$(1-0.8) \times 100 = 20$	$0.8 \times 100 = 80$

Different Set-Up

Now say for the same machine H_A is true 40% of the time. i.e. $P(H_A) = 0.4$

Different Set-Up

Now say for the same machine H_A is true 40% of the time. i.e. $P(H_A) = 0.4$

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	$(1 - 0.05) \times 600 = 570$	$0.05 \times 600 = 30$
	H_A true	$(1 - 0.8) \times 400 = 80$	$0.8 \times 400 = 320$

How Reliable Are Your Test Results?

How Reliable Are Your Test Results?

The probability that a positive result is right depends on how likely H_A is. Same goes for negative results.

How Reliable Are Your Test Results?

The probability that a positive result is right depends on how likely H_A is. Same goes for negative results.

Question 2 from Quiz 9: (Nature article) Say a scientist obtains a p-value of 0.01. An incorrect interpretation of this is that it is the probability of a “false alarm” (type I error)... If one wants to make a statement about this being a false alarm, what additional piece of information is required?

How Reliable Are Your Test Results?

The probability that a positive result is right depends on how likely H_A is. Same goes for negative results.

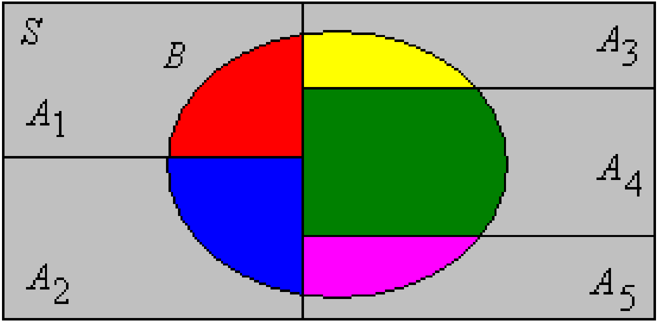
Question 2 from Quiz 9: (Nature article) Say a scientist obtains a p-value of 0.01. An incorrect interpretation of this is that it is the probability of a "false alarm" (type I error)... If one wants to make a statement about this being a false alarm, what additional piece of information is required?

Answer 2: The plausibility of the hypothesis being tested for.

Bayes Theorem

Illustration

- The sample sample S is the overall grey box
- $A_1, \ldots, A_5$ are the five blocks that partition S.
- B is the oval



Tailored to our Situation

Tailored to our Situation

The Debate

Previously, you applied Bayes Theorem from scratch/intuition.

The Debate

Previously, you applied Bayes Theorem from scratch/intuition.

Why isn't everybody taking into account $P(H_A)$ when testing H_0 vs H_A?

The Debate

Previously, you applied Bayes Theorem from scratch/intuition.

Why isn't everybody taking into account $P(H_A)$ when testing H_0 vs H_A?

In this example, we assumed we knew the true $P(H_A)$. In real life however, we don't.

Statistics In General

Statistics is inferring about some unknown parameter θ.

Statistics In General

Statistics is inferring about some unknown parameter θ.

- Frequentist Statistics: the true θ is a single value.

Statistics In General

Statistics is inferring about some unknown parameter θ.

- Frequentist Statistics: the true θ is a single value.
- Bayesian Statistics: the true θ is a distribution of values that reflects our belief in the plausibility of different values.

Statistics In General

Statistics is inferring about some unknown parameter θ.

- Frequentist Statistics: the true θ is a single value.
- Bayesian Statistics: the true θ is a distribution of values that reflects our belief in the plausibility of different values.

Ex: Coin Flips

The Bayesian Procedure

To express our belief about θ from as a Bayesian, we have:

The Bayesian Procedure

To express our belief about θ from as a Bayesian, we have:

1. A prior distribution $\Pr(\theta)$. It reflects our prior belief about θ.

The Bayesian Procedure

To express our belief about θ from as a Bayesian, we have:

1. A prior distribution $\Pr(\theta)$. It reflects our prior belief about θ.
2. The likelihood function $\Pr(X|\theta)$. This is the mechanism that generates the data.

The Bayesian Procedure

To express our belief about θ from as a Bayesian, we have:

1. A prior distribution $\Pr(\theta)$. It reflects our prior belief about θ.
2. The likelihood function $\Pr(X|\theta)$. This is the mechanism that generates the data.
3. A posterior distribution $\Pr(\theta|X)$. We update our belief about θ after observing data X.

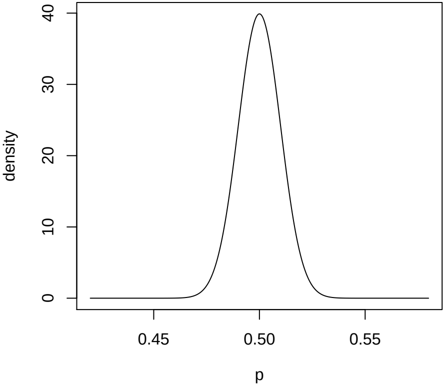
$$\Pr(\theta|X) = \frac{\Pr(X|\theta) \cdot \Pr(\theta)}{\Pr(X)}$$

The Issue: The Bayesian Procedure

Where do you come up with $\Pr(\theta)$? It's completely subjective! You decide!

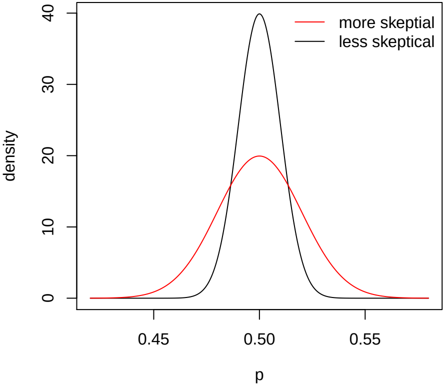
Prior Distribution

This distribution can reflect someone's prior belief of p.



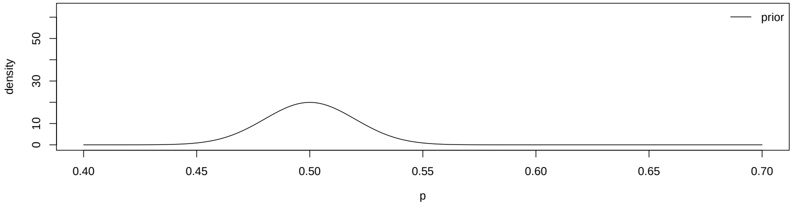
Prior Distribution

Say someone is more skeptical that $p = 0.5$, we can lower it.



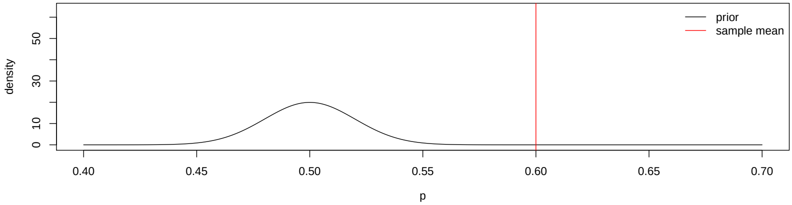
The Bayesian Procedure

Say we have the following prior belief centered at $p = 0.5$



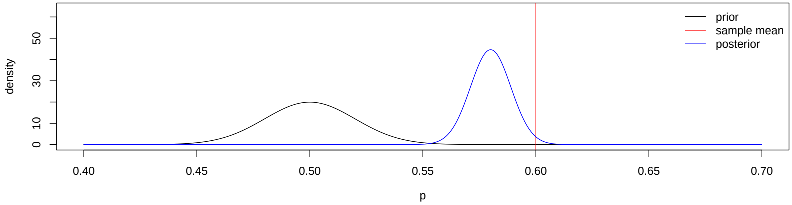
The Bayesian Procedure

Say we collect data, represented by the red line, suggesting $p = 0.6$



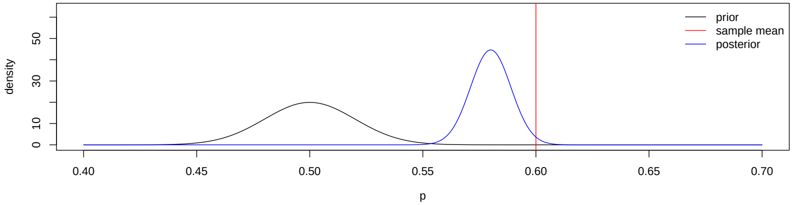
The Bayesian Procedure

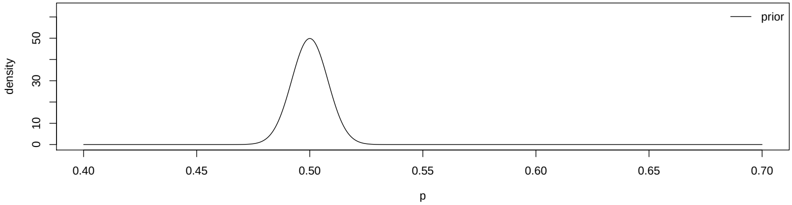
We then update our belief, as reflected in the posterior distribution!



The Bayesian Procedure

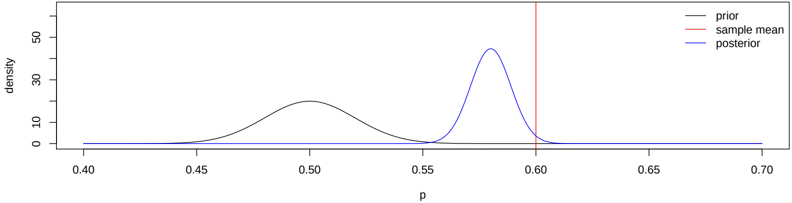
Now say we have a stronger prior belief that $p = 0.5$

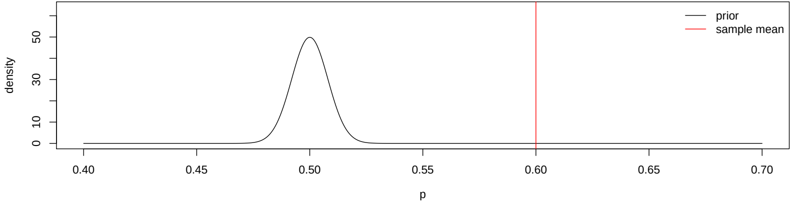




The Bayesian Procedure

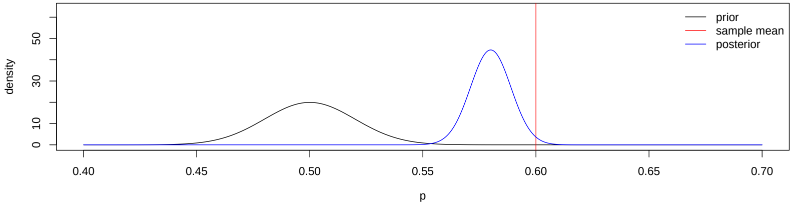
Say we observed the same data (as represented in red).

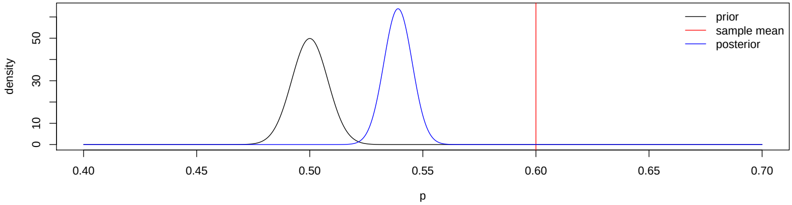




The Bayesian Procedure

The posterior in this case is pulled left due to the sharper prior.





Back to Debate

Frequentists believe statistics should be completely objective and therefore do not accept the premise of a subjective prior.

Back to Debate

Frequentists believe statistics should be completely objective and therefore do not accept the premise of a subjective prior.

Back to Hypothesis Testing Machine example. In order to use such a procedure in real life, we would need to specify some prior belief $\Pr(H_A)$ that H_A is true.