* Julia is having an outdoor wedding ceremony tomorrow. In recent years, it has rained on average 50 days/year. Unfortunately, a meteorologist has predicted rain for her wedding day. When it rains, the meteorologist will have correctly predicted it 80% of the time. When it does not, the meteorologist will have incorrectly predicted rain 30% of the time. Given this info, what is the probability that it rains on Julia's wedding day?
* P(R | predicted) = P(R and predicted C) / P(predicted C) =

> p\_rain <- 50/365

> p\_no\_rain <- 1 - p\_rain

>

> p\_correct\_given\_rain <- .8

> p\_correct\_given\_no\_rain <- .3

>

> p\_rain\_and\_correct <- p\_rain\*p\_correct\_given\_rain

> p\_no\_rain\_and\_correct <- p\_no\_rain\*p\_correct\_given\_no\_rain

> p\_correct <-p\_rain\_and\_correct + p\_no\_rain\_and\_correct

>

> (p\_rain\_giv\_pred <- p\_rain\_and\_correct\_rain / p\_correct)

[1] 0.2973978

* **29.7%**
* Which of the following do impacts decisions based on frequentist inference?
* ~~Posterior probability~~
* **Significance level**
* **Type 1 error rate**
* **The null hypothesis**
* Suppose 20 people are randomly sampled from the population + their gender is recorded. Which of the following best represents the **likelihood** of the number of males observed k?
* **The probability of observing exactly k males in 20 people, given p (the true population proportion of males)**
* Which of the following is consistent w/ both Bayesian + frequentist interpretations of probability?
* ~~Probability can be represented by long-run frequency of an event divided by # of trials.~~ (F)
* ~~Probability is the tendency of an experiment to produce a certain outcome, even if it is performed only once.~~
* **Probability is a measure of the likelihood that an event will occur.**
* ~~Probability can be represented by a degree of belief, which changes as w/ data collected.~~ (B)
* You are told a coin has 1 of the following, w/ probability of heads under that event noted next to it in parentheses:
* a strong tails bias (p = 0.2)
* a weak tails bias (p = 0.4)
* no bias (p = 0.5)
* a weak heads bias (p = 0.6)
* a strong heads bias (p = 0.8)
* You assign a prior probability of 1/2 that the coin is fair + distribute the remaining 1/2 prior probability equally over the other 4 possible scenarios (.5/4). You flip the coin 3 times + it comes up heads all 3 times. What is the posterior probability that the coin is biased towards heads?

> k <- 3

> n <- 3

> p\_values <- c(.2,.4,.5,.6,.8)

> priors <- c(rep(.5/4,2),.5,rep(.5/4,2))

> (likelihood <- dbinom(k, n, prob = p\_values))

[1] 0.008 0.064 0.125 0.216 0.512

> (posteriors <- (priors \* likelihood) / sum((priors \* likelihood)))

[1] 0.006153846 0.049230769 0.384615385 0.166153846

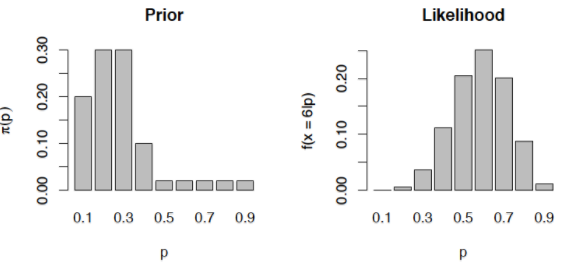
[5] 0.393846154

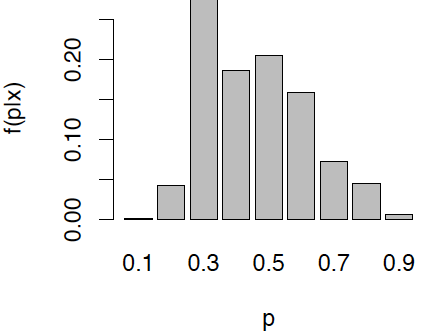
> sum(posteriors) # check they sum to 1

[1] 1

> sum(posteriors[4:5])

[1] 0.56

* **0.56**
* You draw 2 balls from 1 of 3 possible urns, labelled A, B, C. Urn A has 1/2 blue balls, 1/3 green balls, and 1/6 red balls. Urn B has 1/6 blue balls, 1/2 green balls, and 1/3 red balls. Urn C has 1/3 blue balls, 1/6 green balls, and 1/2 red balls. W/ no prior info about which urn you’re drawing from, you draw 1 red ball + 1 blue ball. What is the probability that you drew from urn C?
* 19/36
* 5/9
* 6/11
* 1/3
* Below are plots of the prior distribution for a success probability p + the likelihood as a function of p, where 6 successes were observed in ten trials. Which of the following is most likely to be the posterior distribution for the proportion p?
* 



* You go to Las Vegas + sit down at a slot machine + are told by a highly reliable source that, for each spin, the probability of hitting the jackpot is either 1 in 1,000 or 1 in 1,000,000, but you have no prior info to tell you which it is. You play 10 times, but do not win the jackpot. What is the posterior probability that the true odds of hitting the jackpot are 1 in 1,000?

> k <- 0

> n <- 10

> p\_values <- c(1/1000,1/1000000)

> priors <- c(.5,.5)

> (likelihood <- dbinom(k, n, prob = p\_values))

[1] 0.9900449 0.9999900

> (posteriors <- (priors \* likelihood) / sum((priors \* likelihood)))

[1] 0.4975013 0.5024987

* **0.498**
* The posterior distribution after repeating the same experiment twice + analyzing the data from both experiments at the same time is the same as after running the 2nd experiment w/ the posterior of the 1st experiment as the prior.
* **False**
* Which of the following is the best *Bayesian* interpretation of the following statement: ”The probability of Liverpool defeating Swansea City tomorrow is 0.9“?
* ~~Liverpool would beat Swansea City nine times out of ten~~ (F)
* **We would be indifferent to betting on Liverpool to win at 1:9 odds.**
* Teams as good as Liverpool have historically beaten teams as good as Swansea City 90% of the time.
* Liverpool is a heavy favorite to beat Swansea City.
* Which of the following statements can be used to describe a 95% Bayesian *credible* interval for a parameter μ, but not a 95% Frequentist CI?
* μ is in this interval 95 percent of the time.
* If we ran an infinite number of experiments, 95 percent of our intervals generated this way would contain the true value of μ .
* **μ is either in the interval, or it is not. More data can increase or decrease our uncertainty that μ is in the interval.**
* The probability that μ falls within the interval is 0.95
* A new breast cancer screening method is tested to see if it performs better than existing methods. To measure **sensitivity** of the test, a total of 10k patients known to have various stages of breast cancer are testing using the new method. Of those 10k patients, 9,942 are identified by the new method to have breast cancer. Given that **sensitivity** of the best current test = 99.3%, is there significant evidence at the α = 0.05 level to conclude the new method has higher sensitivity than existing methods? Hint - H0: p = 0.993 and H1: p > 0.993

> binom.test(x=9942, n=10000, p=.993, alternative="greater")

Exact binomial test

data: 9942 and 10000

number of successes = 9942, number of trials =

10000, p-value = 0.08086

alternative hypothesis: true probability of success is greater than 0.993

95 percent confidence interval:

0.9927874 1.0000000

sample estimates:

probability of success

0.9942

* **No, since the p-value under H0 of no difference is approximately equal to 0.081, which is greater than α = 0.05**
* In the NFL, there are 32 teams, of which 12 make the playoffs. In a typical season, 20 teams (ones that don’t make the playoffs) play 16 games, 4 teams play 17 games, 6 teams play 18 games, + 2 teams play 19 games. At the beginning of each game, a coin is flipped to determine who gets the football 1st. You are told an unknown team won 10 of its coin flips last season. Given this info, what is the posterior probability that the team did not make the playoffs (i.e. played 16 games)?

> k <- 10

> n <- 16

> #p\_values <- c(20/32,4/32,6/32,2/32)

> priors\_did\_not\_make\_playoff <- c(20/32,6/32,4/32,2/32)

> (likelihood <- dbinom(k, n, prob = .5))

[1] 0.1221924

> (posteriors <- (priors\_did\_not\_make\_playoff \* likelihood) / sum((priors\_did\_not\_make\_playoff \* likelihood)))

[1] 0.6250 0.1875 0.1250 0.0625

* **0.625**
* You’re testing dice for a casino to make sure 6’s do not come up more frequently than expected. B/c you don’t want to manually roll dice all day, you design a machine to roll a die repeatedly + record the # of 6’s that come up. In order to do a Bayesian analysis to test the hypothesis that p = 1/6 vs. p = .175, you set the machine to roll the die 6k times. You discover to your horror the machine was unable to count higher than 999 + the machine says 999 sixes occurred. Given a prior probability of 0.8 placed on hypothesis p = 1/6 , what is the posterior probability that the die is fair, given the censored data?
* 0.500
* 0.684
* 0.800
* 0.881
* 1
* point
* 10.
* As long as the prior places non-zero probability on all possible values of a proportion, the posterior of the proportion is guaranteed to converge to the true proportion as the sample size approaches infinity.
* **True**
* False
* Upgrade to submit