Bayesian statistics in ecology

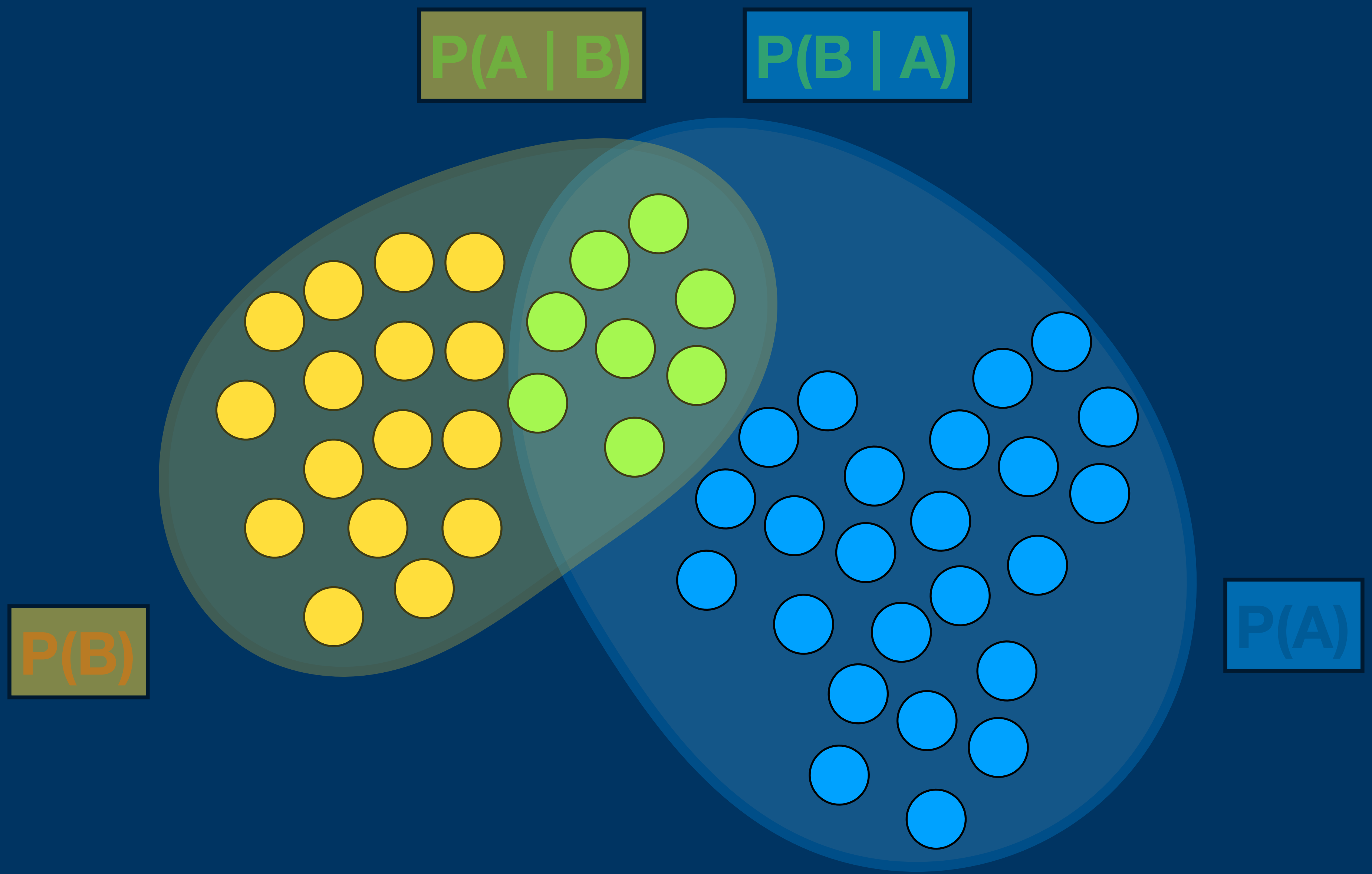
Jérémy Carlot – Postdoc MOGPA 4Oceans

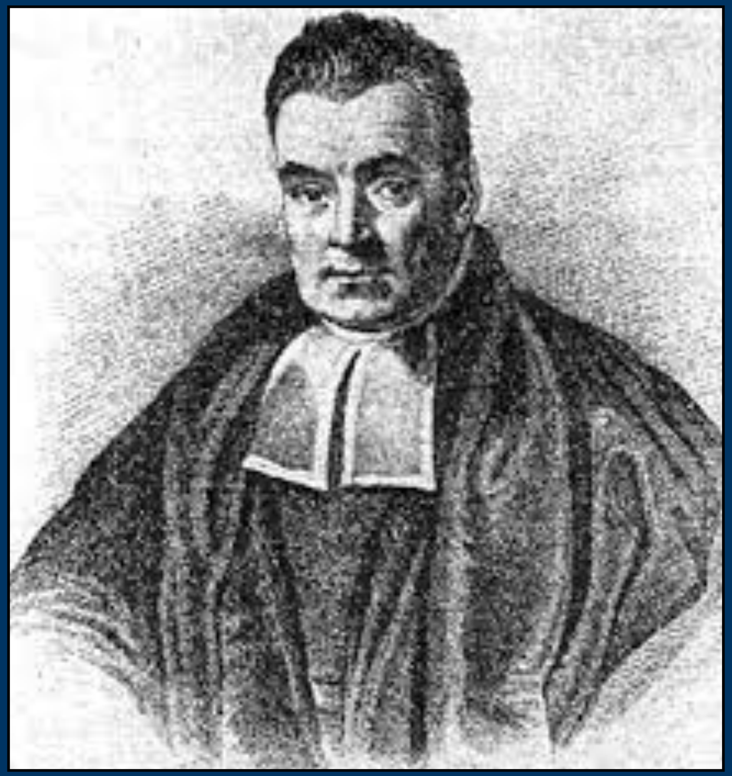
14.02.24

Part 1 – Prerequisites

Bayes theorem

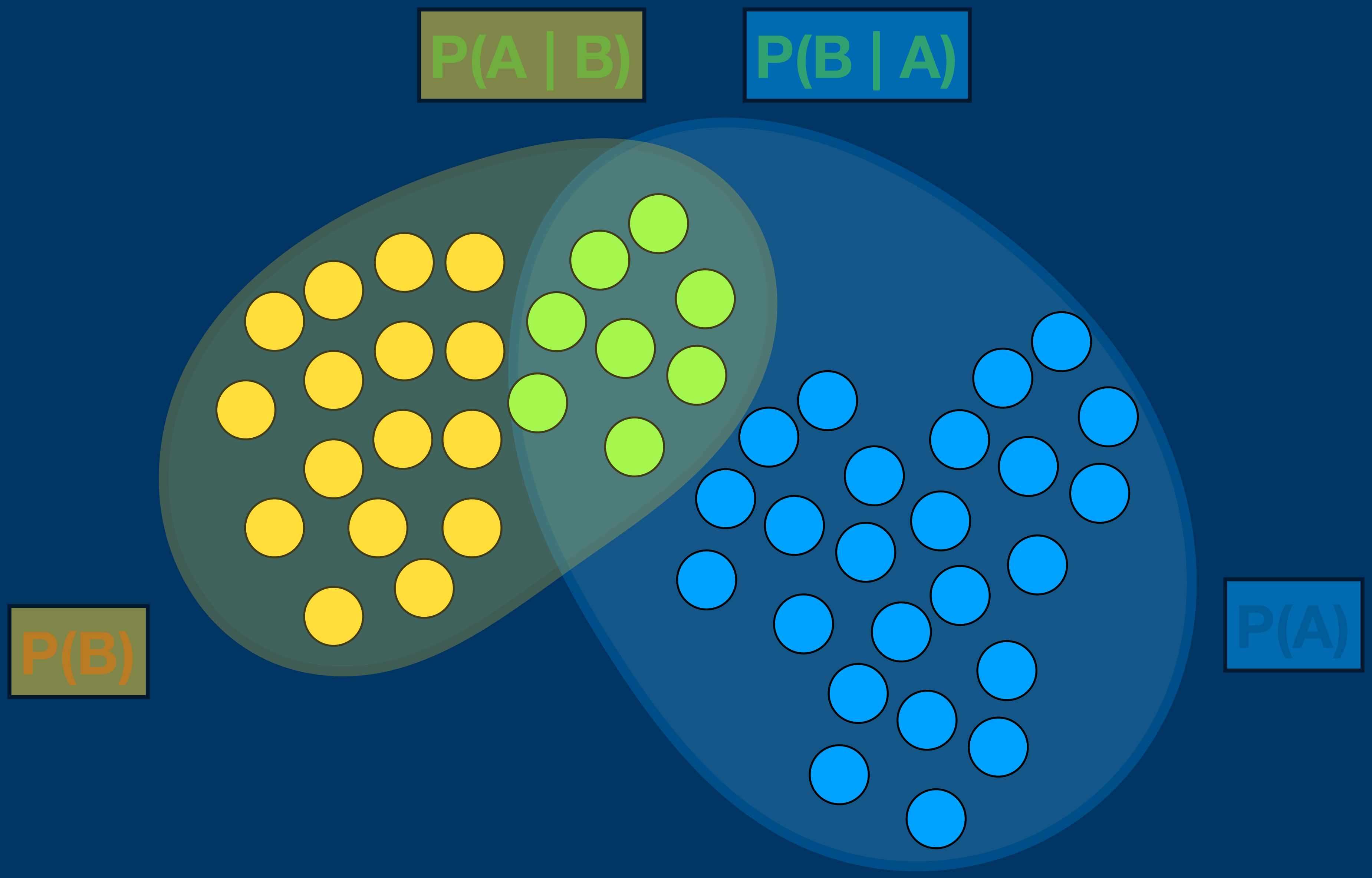
What is Bayesian ?

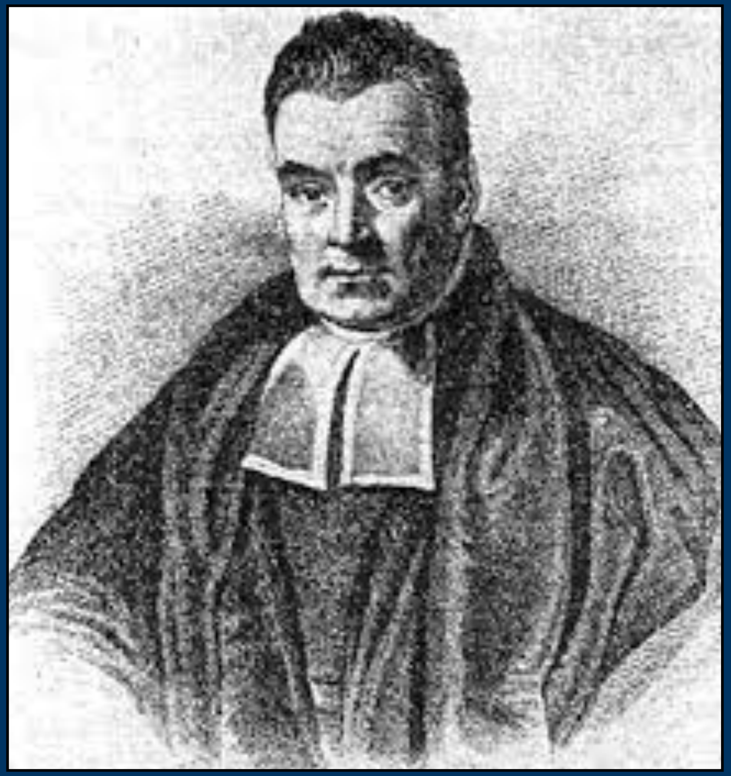




T. Bayes, 1761

What is Bayesian ?

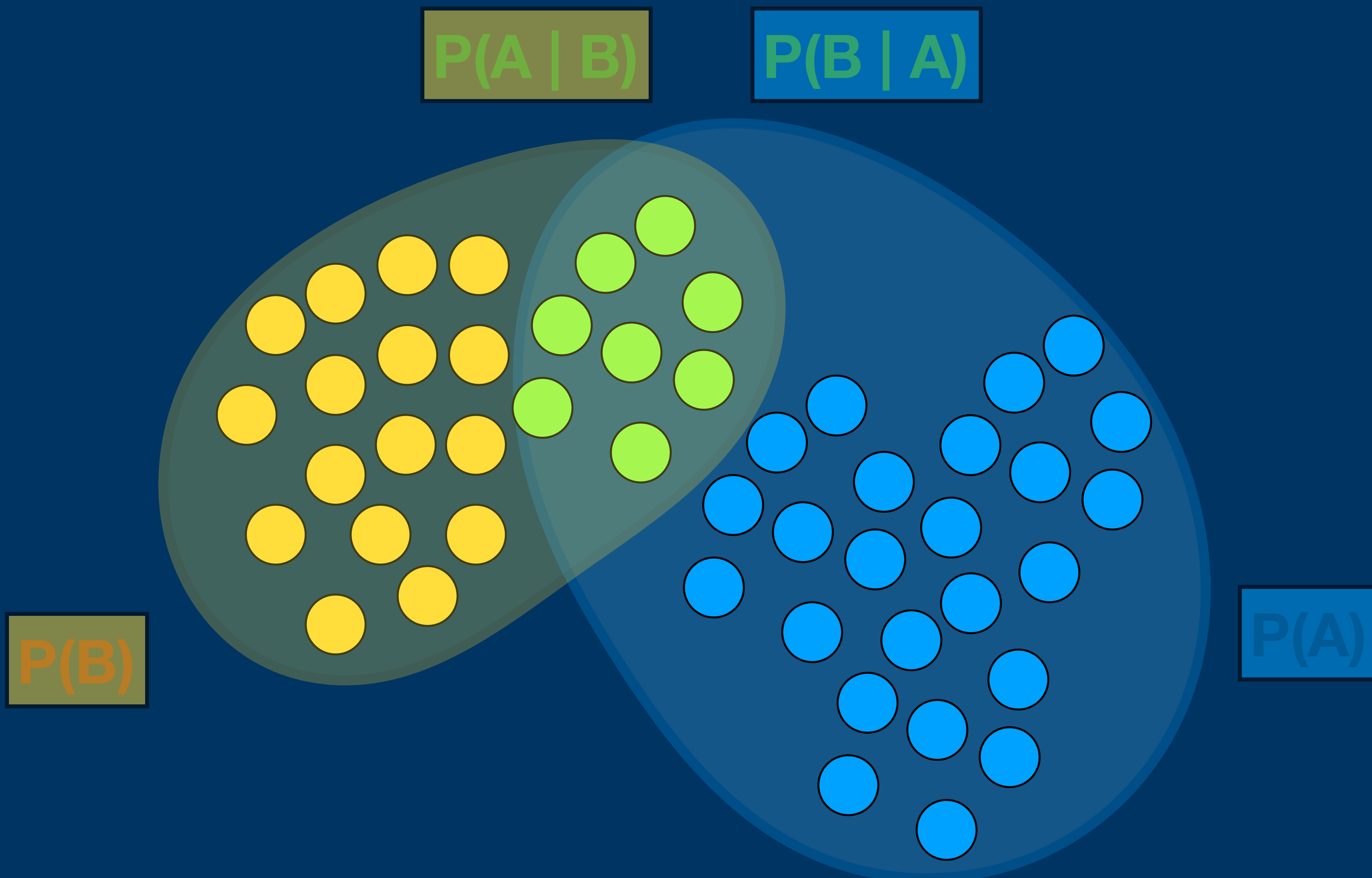


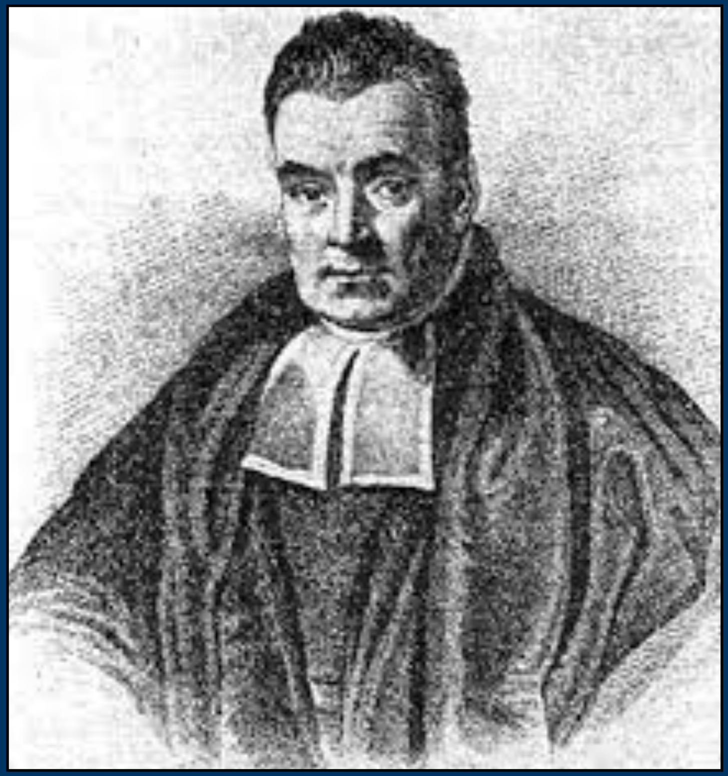


T. Bayes, 1761

$$P(A \cap B) = P(B \mid A) \cdot P(A)$$

What is Bayesian ?



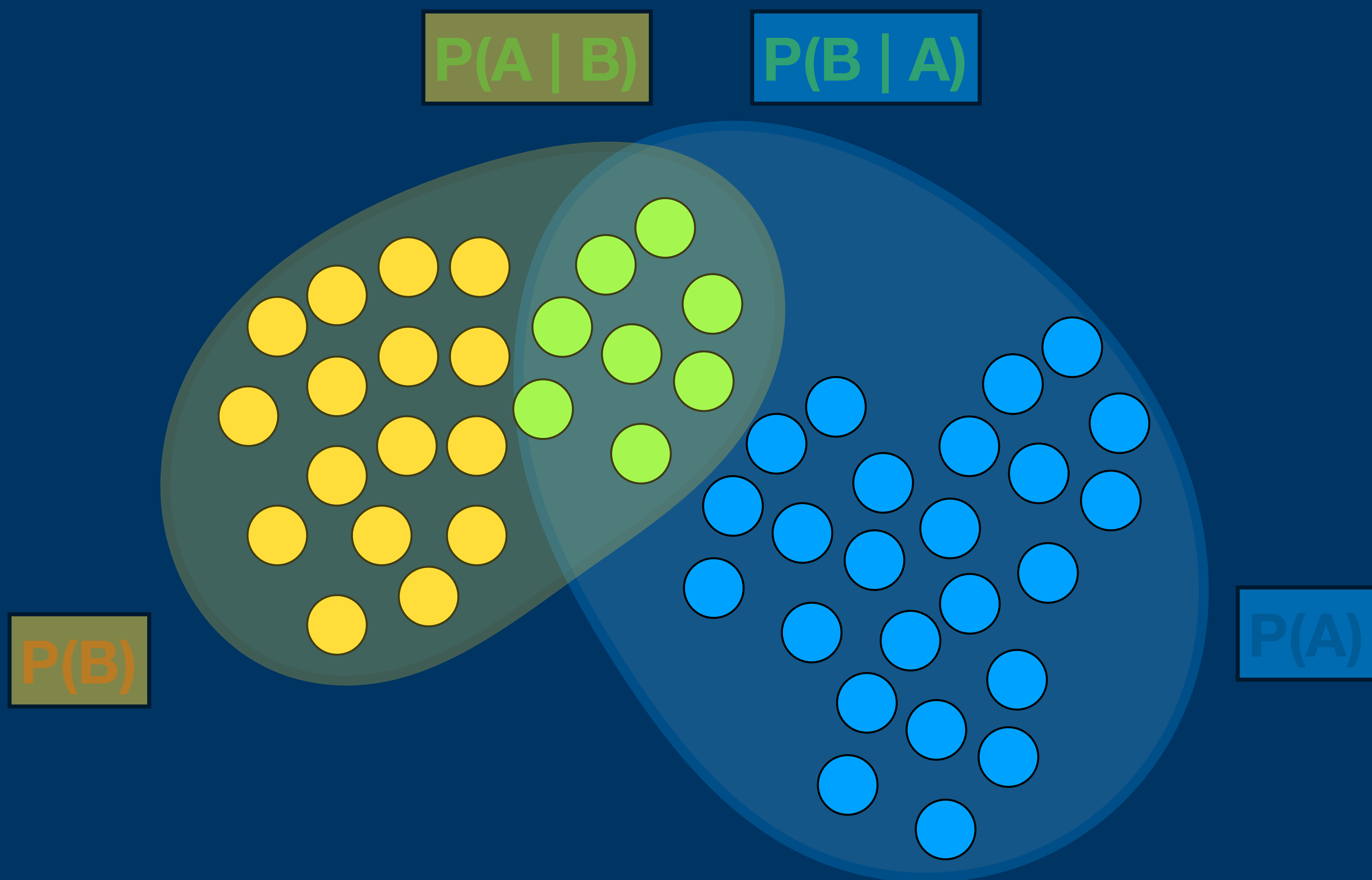


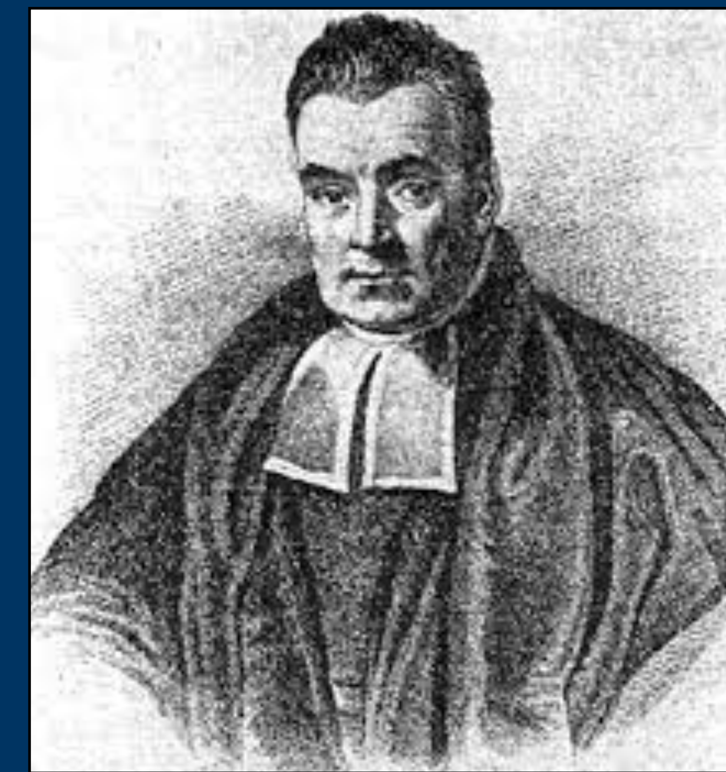
T. Bayes, 1761

$$P(A \cap B) = P(B \mid A) \cdot P(A)$$

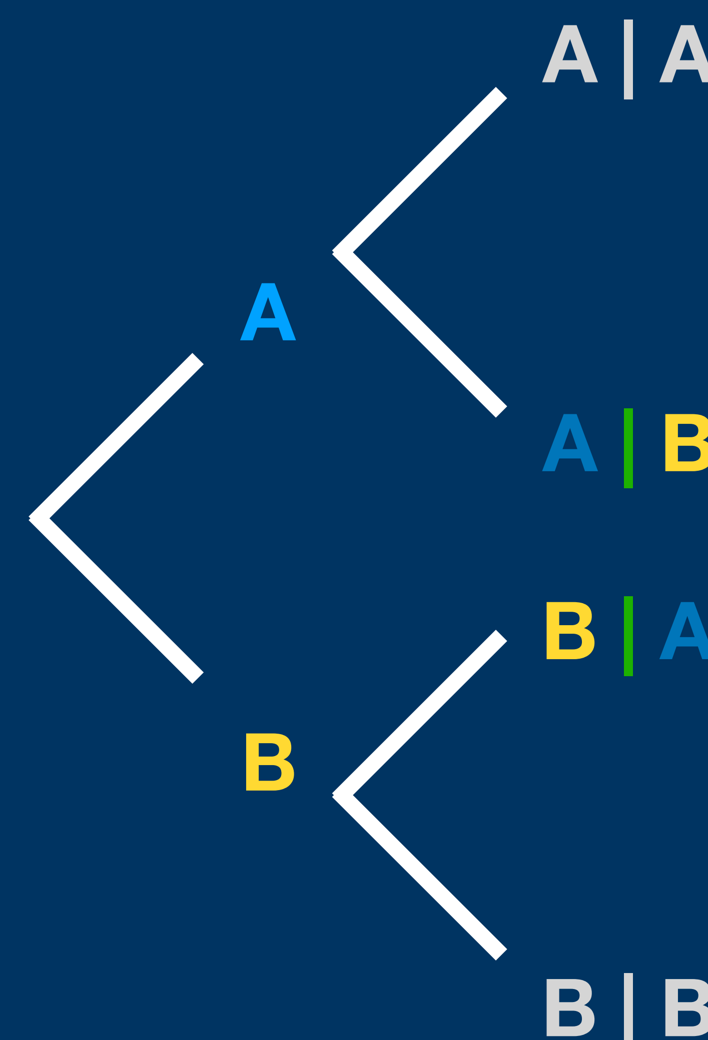
$$P(A \cap B) = P(A \mid B) \cdot P(B)$$

What is Bayesian ?





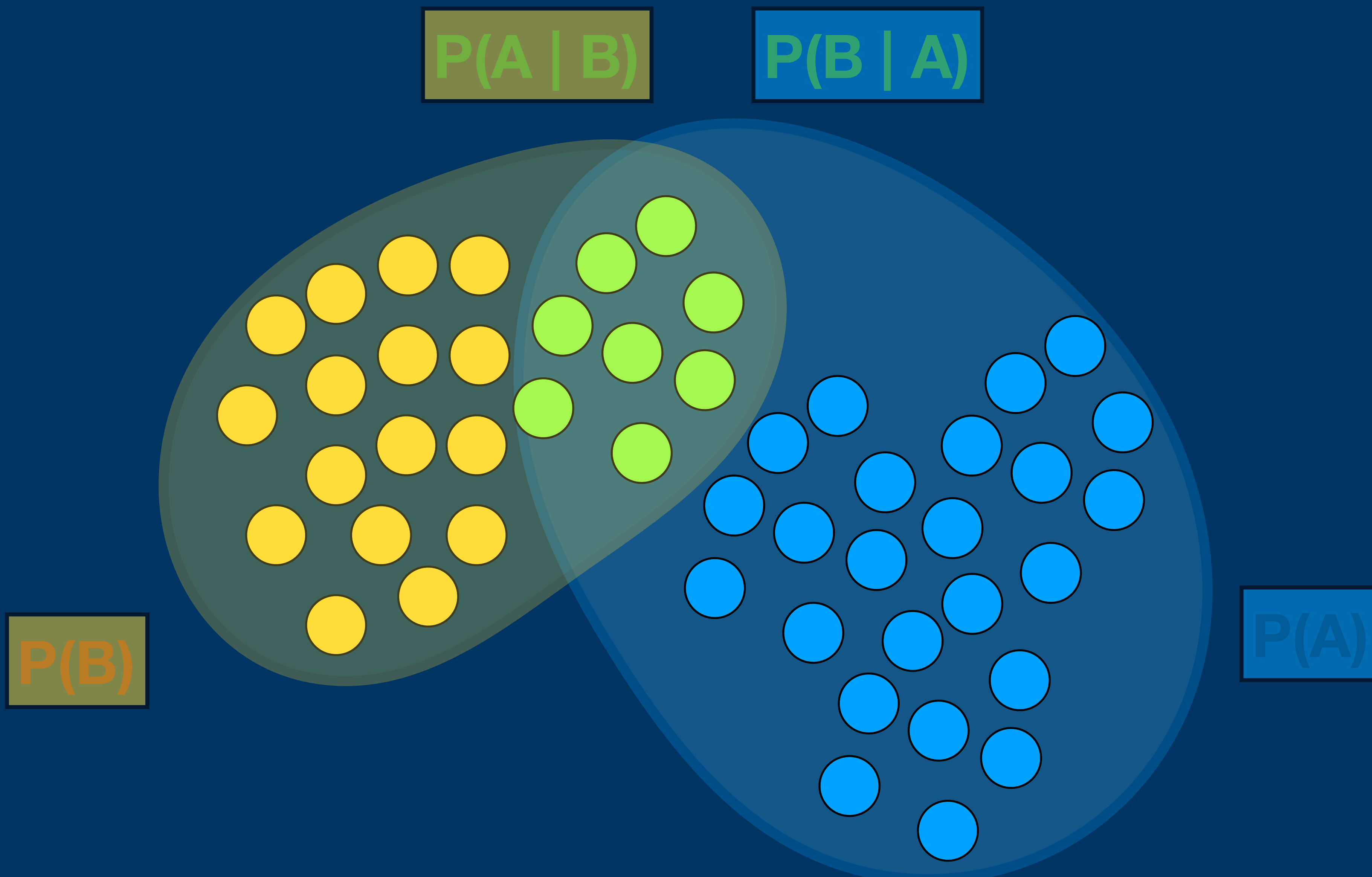
T. Bayes, 1761

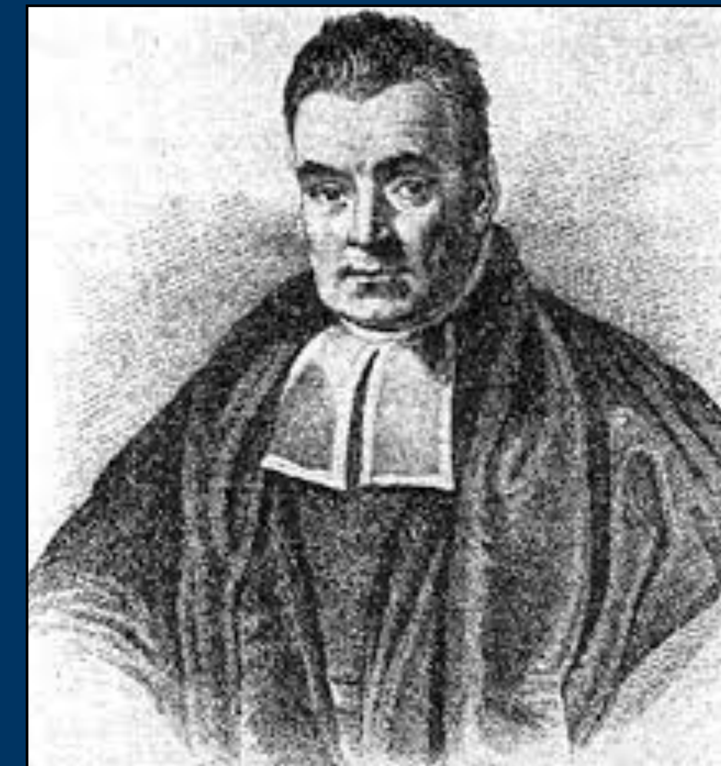


$$P(A \cap B) = P(B \mid A) \cdot P(A)$$

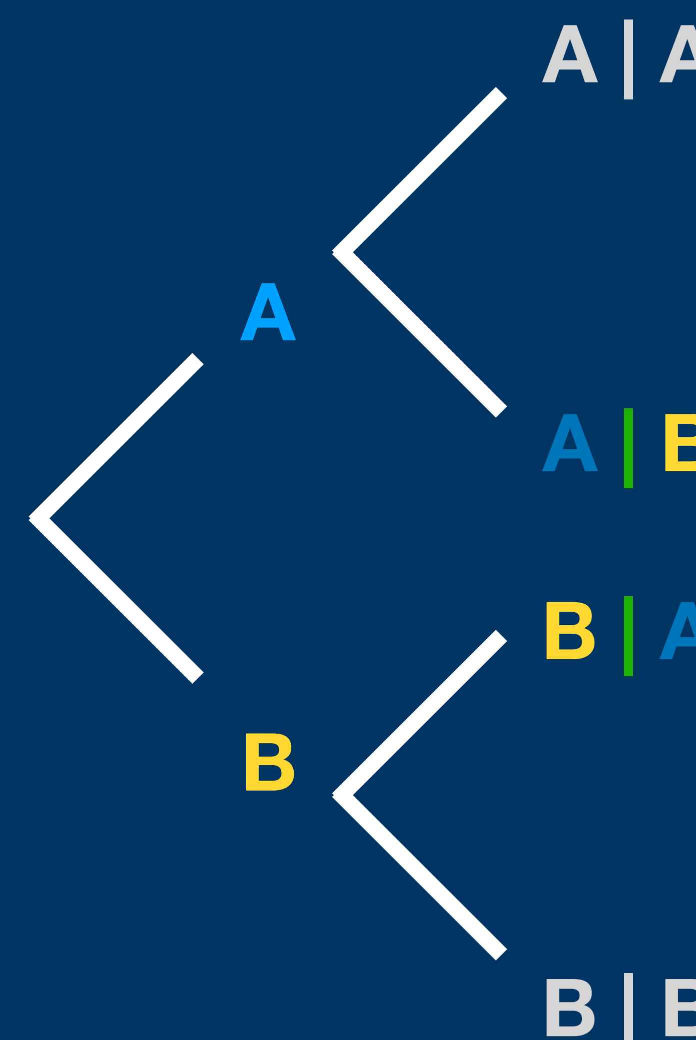
$$P(A \cap B) = P(A \mid B) \cdot P(B)$$

What is Bayesian ?



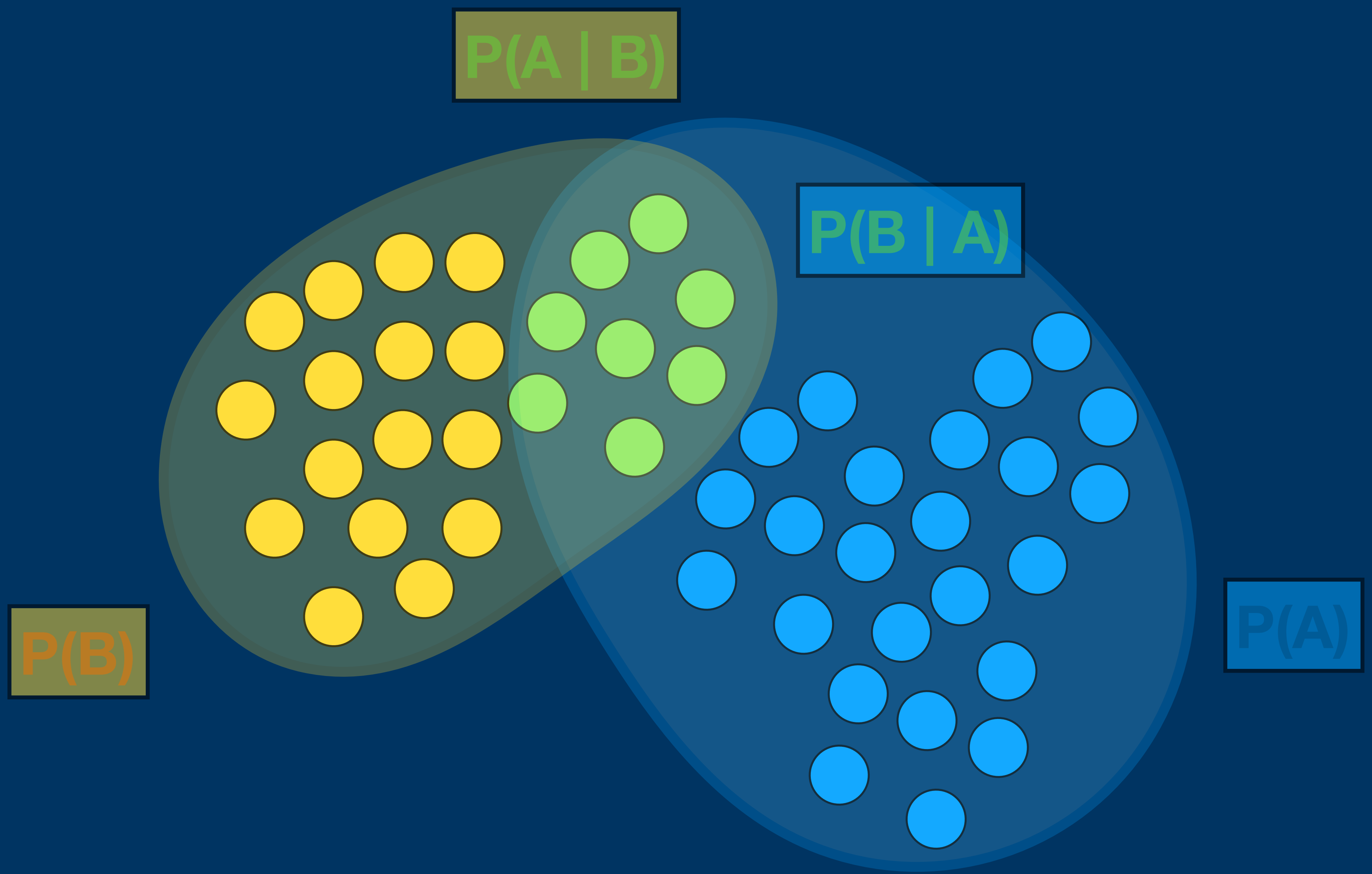


T. Bayes, 1761



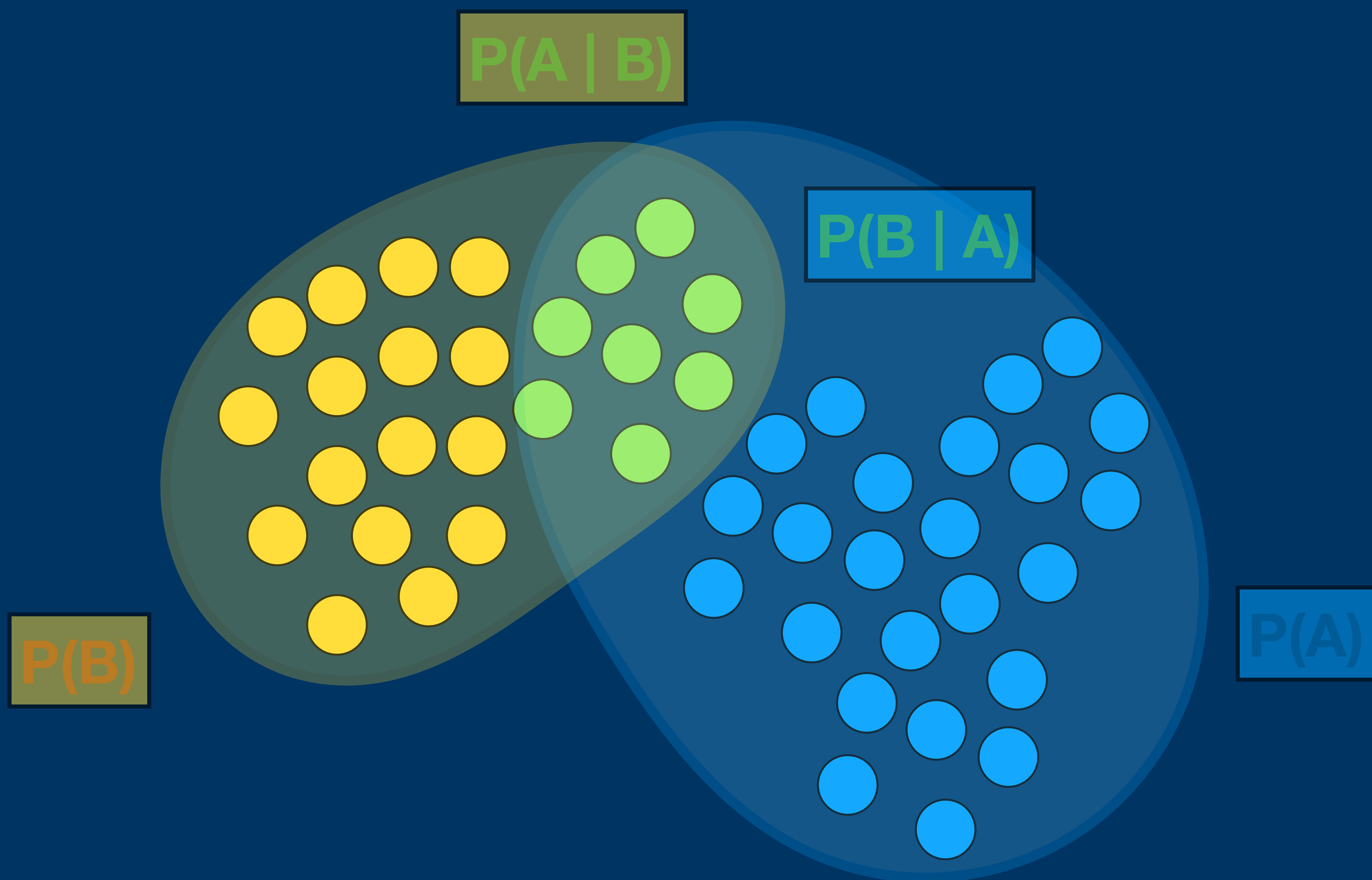
$$P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

What is Bayesian ?



P.-S. Laplace, 1774

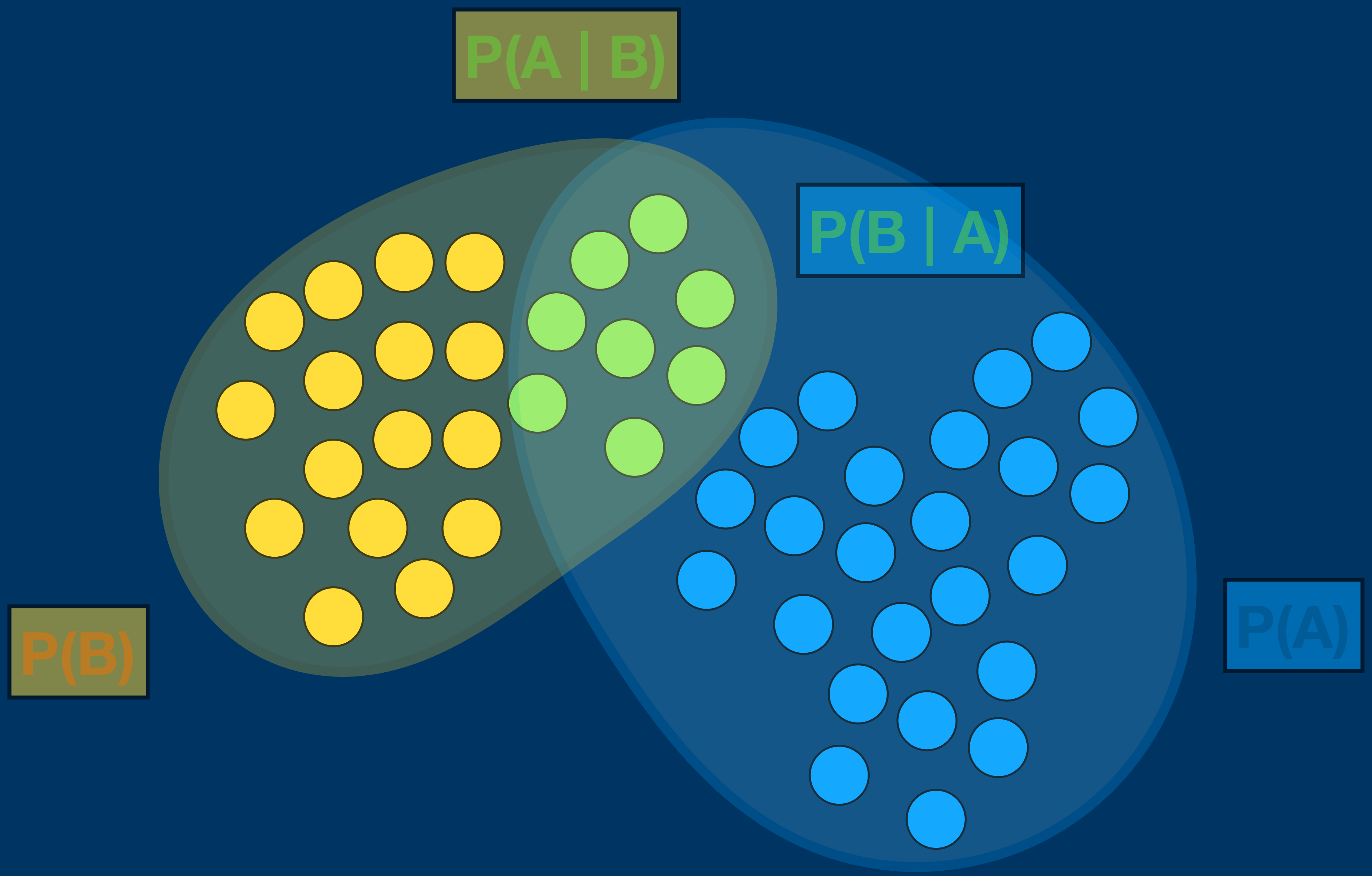
What is Bayesian ?



P.-S. Laplace, 1774

$$P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

What is Bayesian ?



P.-S. Laplace, 1774

$$P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

$$P(B \mid A) = \frac{P(B) \cdot P(A \mid B)}{P(A)}$$

Frequentist *vs.* Bayesian

Frequentist

Likelihood of the results within H_0

$P(x \mid H_0)$

if $P(x \mid H_0) < 0.05$ we reject H_0

R. Fisher, 1930's

Bayesian

Credibility of H_0 within the results

$P(H_0 \mid x)$

–Too complex
–Too subjective

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

Case study – Which dice did I choose ?

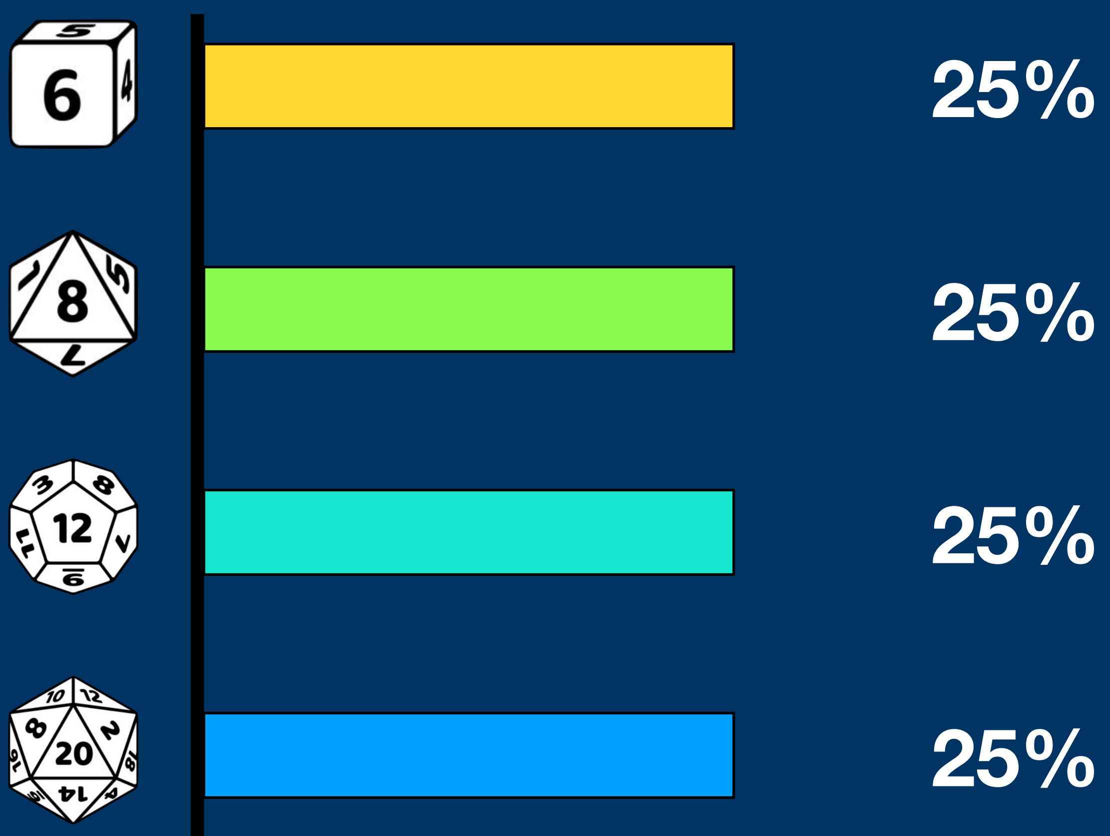
I draw 7

With the frequentist approach

Case study – Which dice did I choose ?

I draw 7

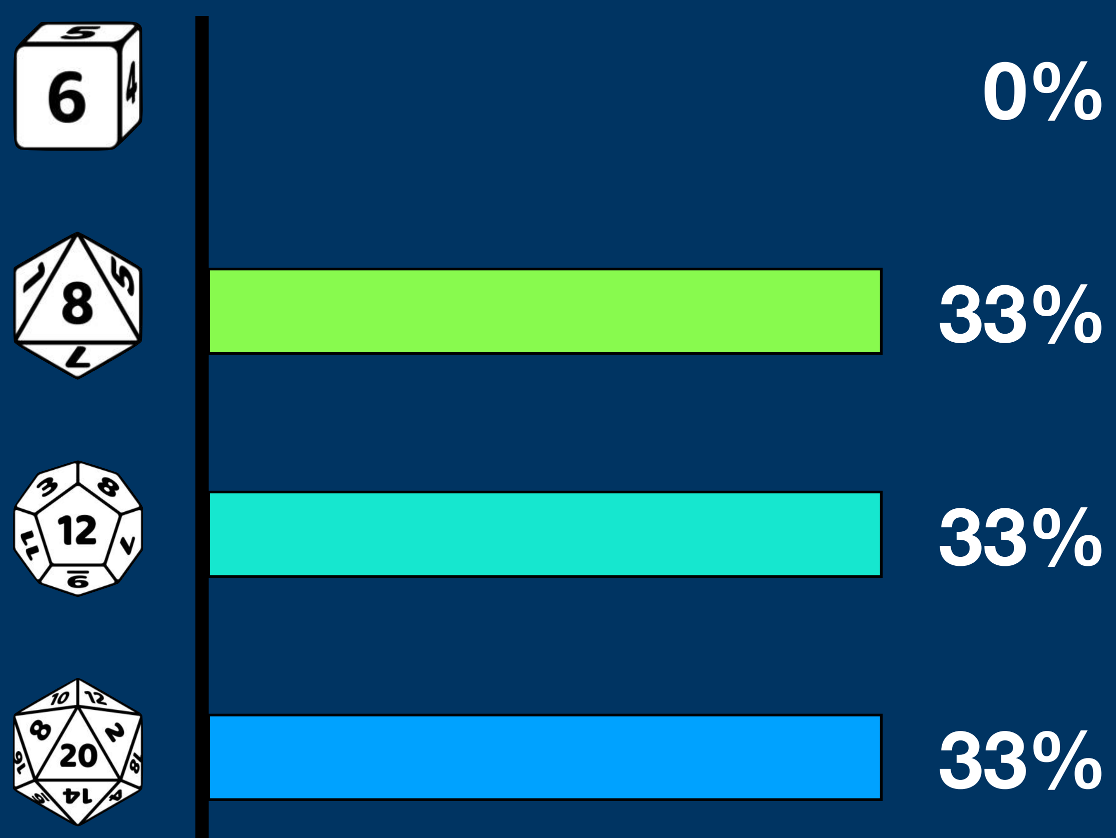
With the frequentist approach



Case study – Which dice did I choose ?

I draw 7

With the frequentist approach

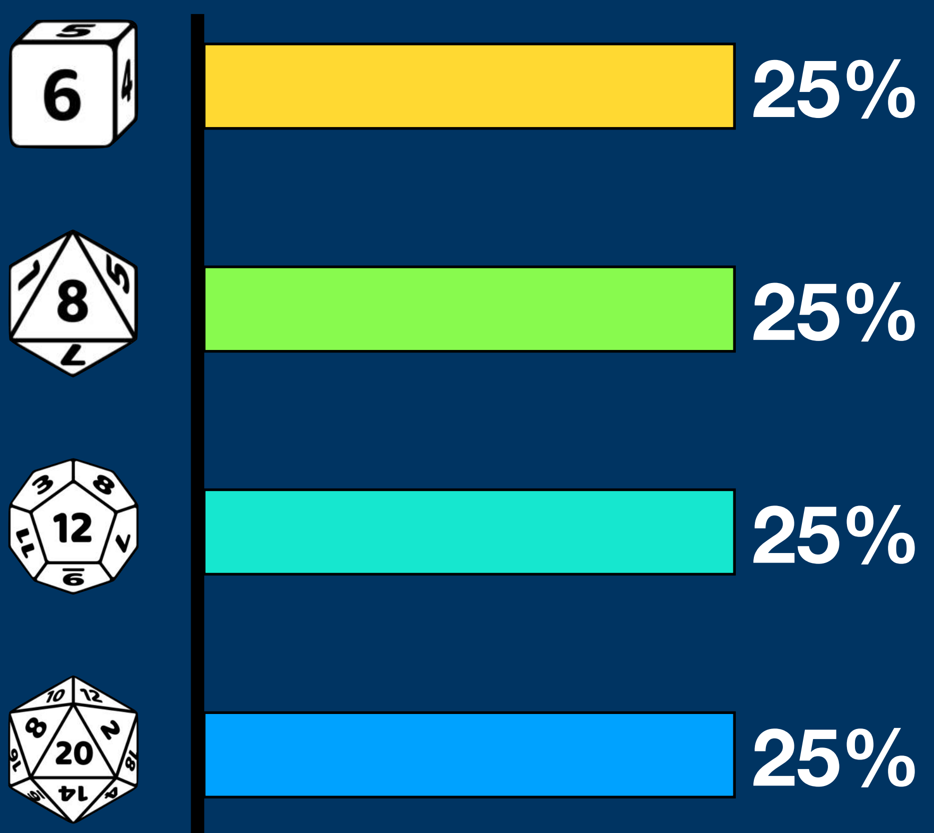


But then, you’re stuck…

Case study – Which dice did I choose ?

I draw 7

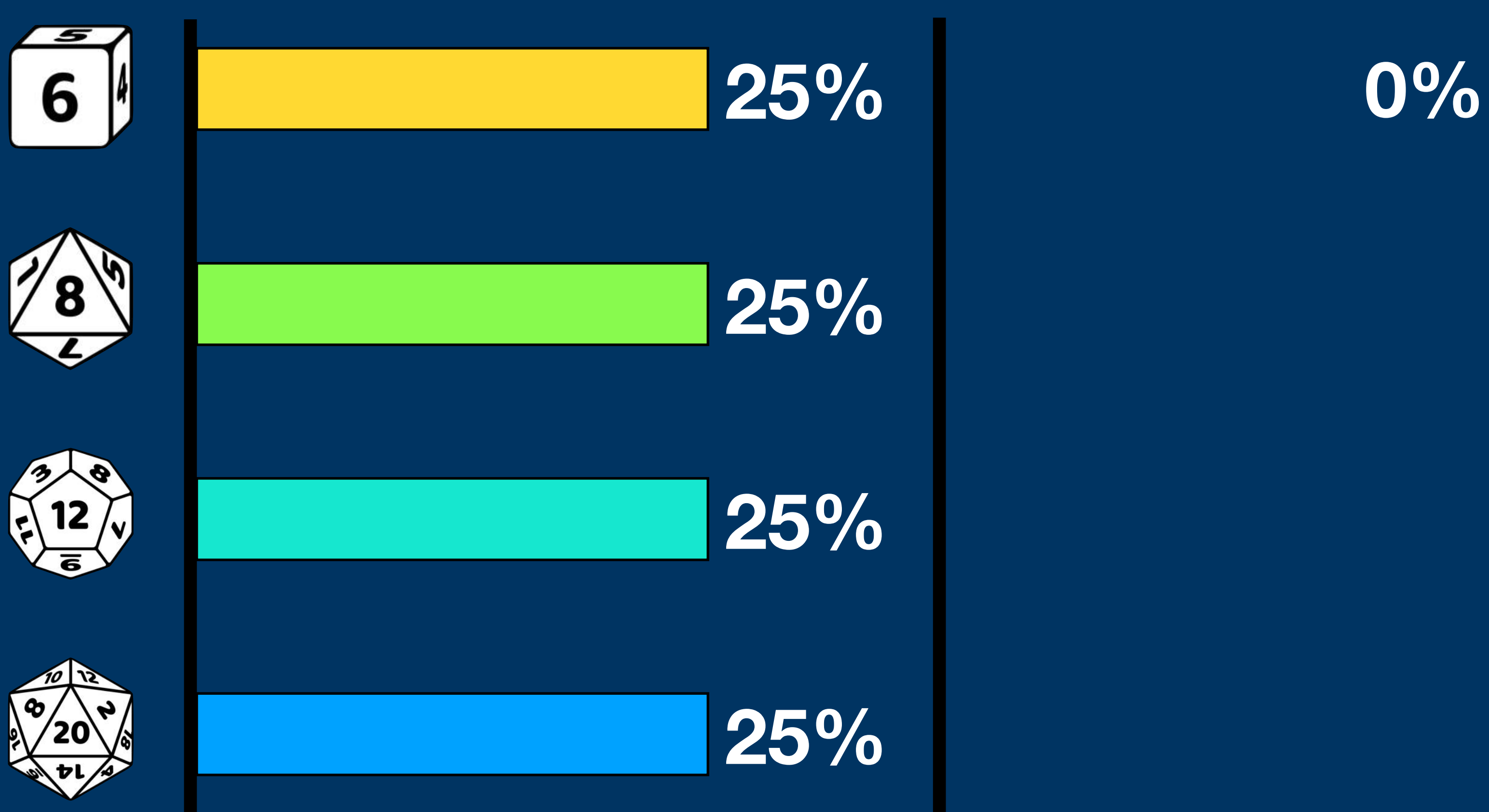
With the bayesian approach



Case study – Which dice did I choose ?

I draw 7

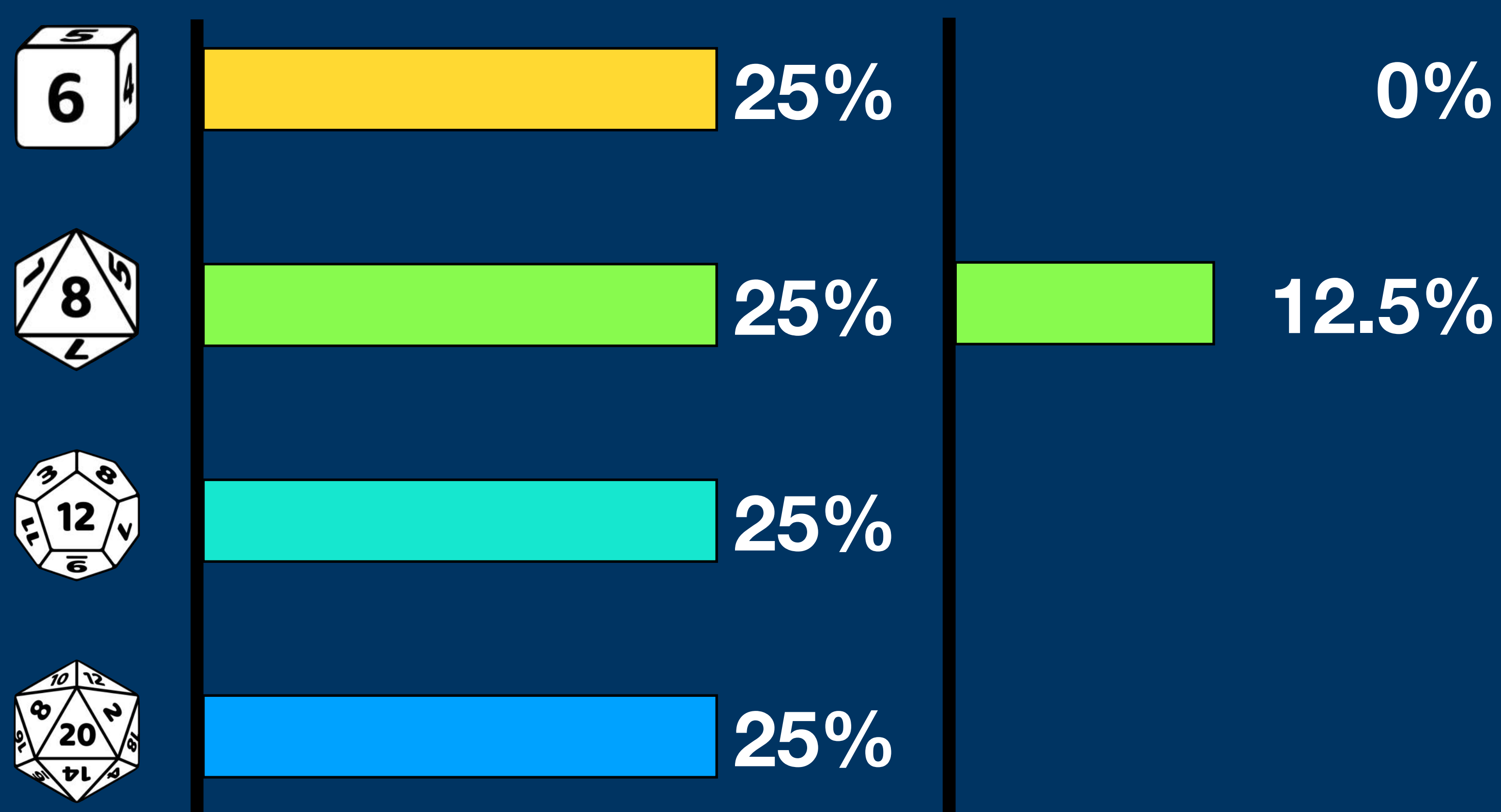
With the bayesian approach



Case study – Which dice did I choose ?

I draw 7

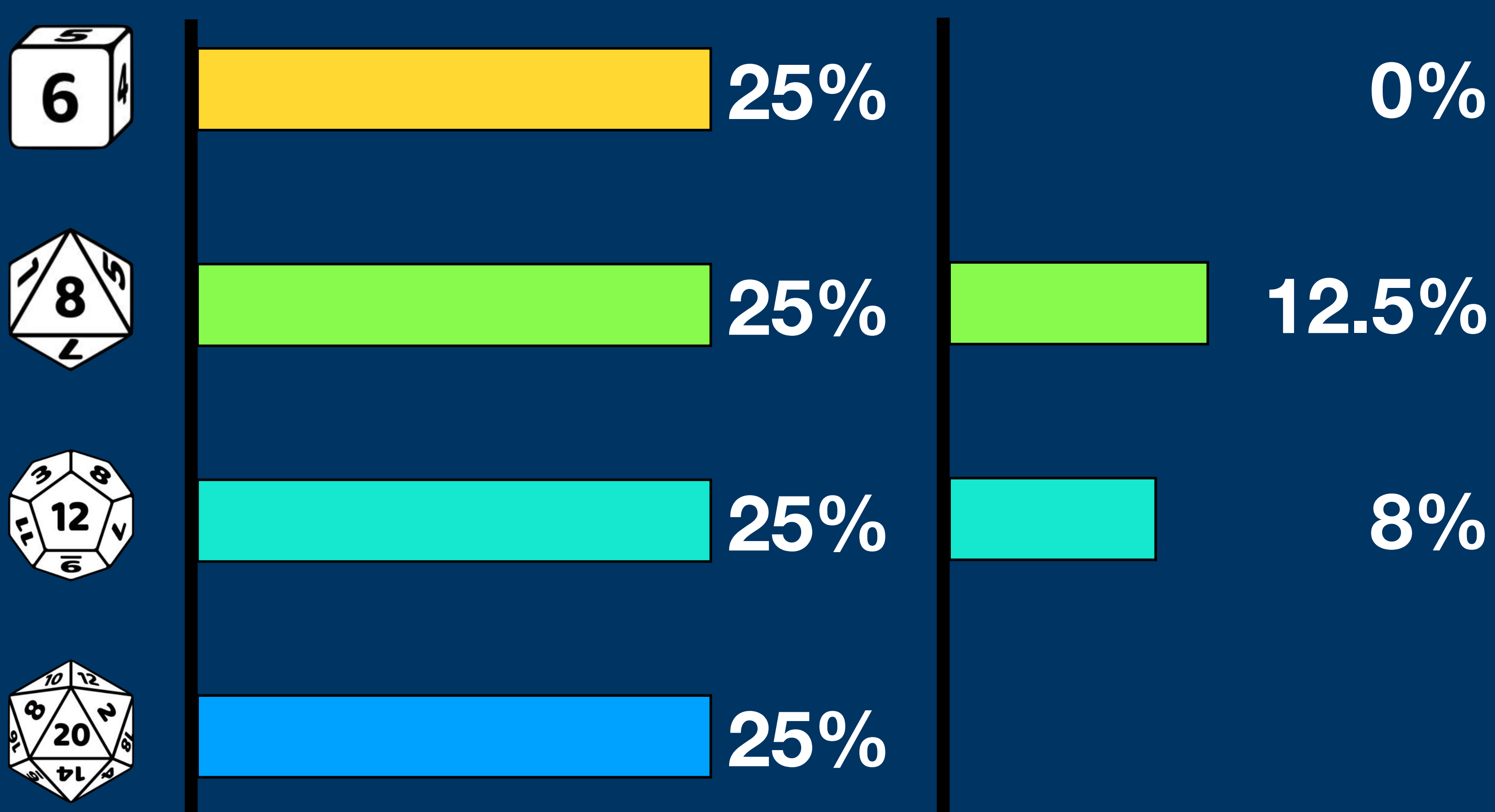
With the bayesian approach



Case study – Which dice did I choose ?

I draw 7

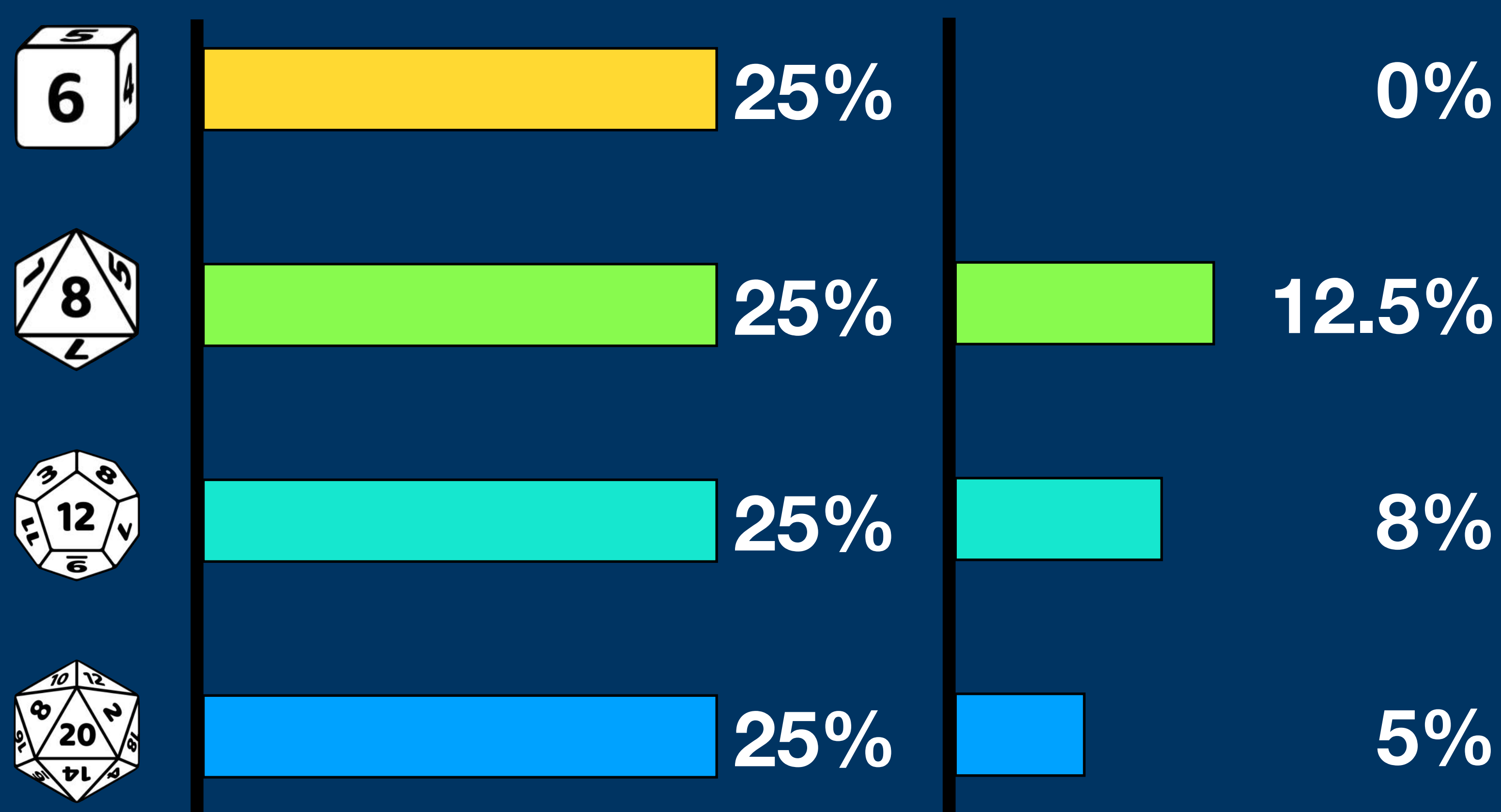
With the bayesian approach



Case study – Which dice did I choose ?

I draw 7

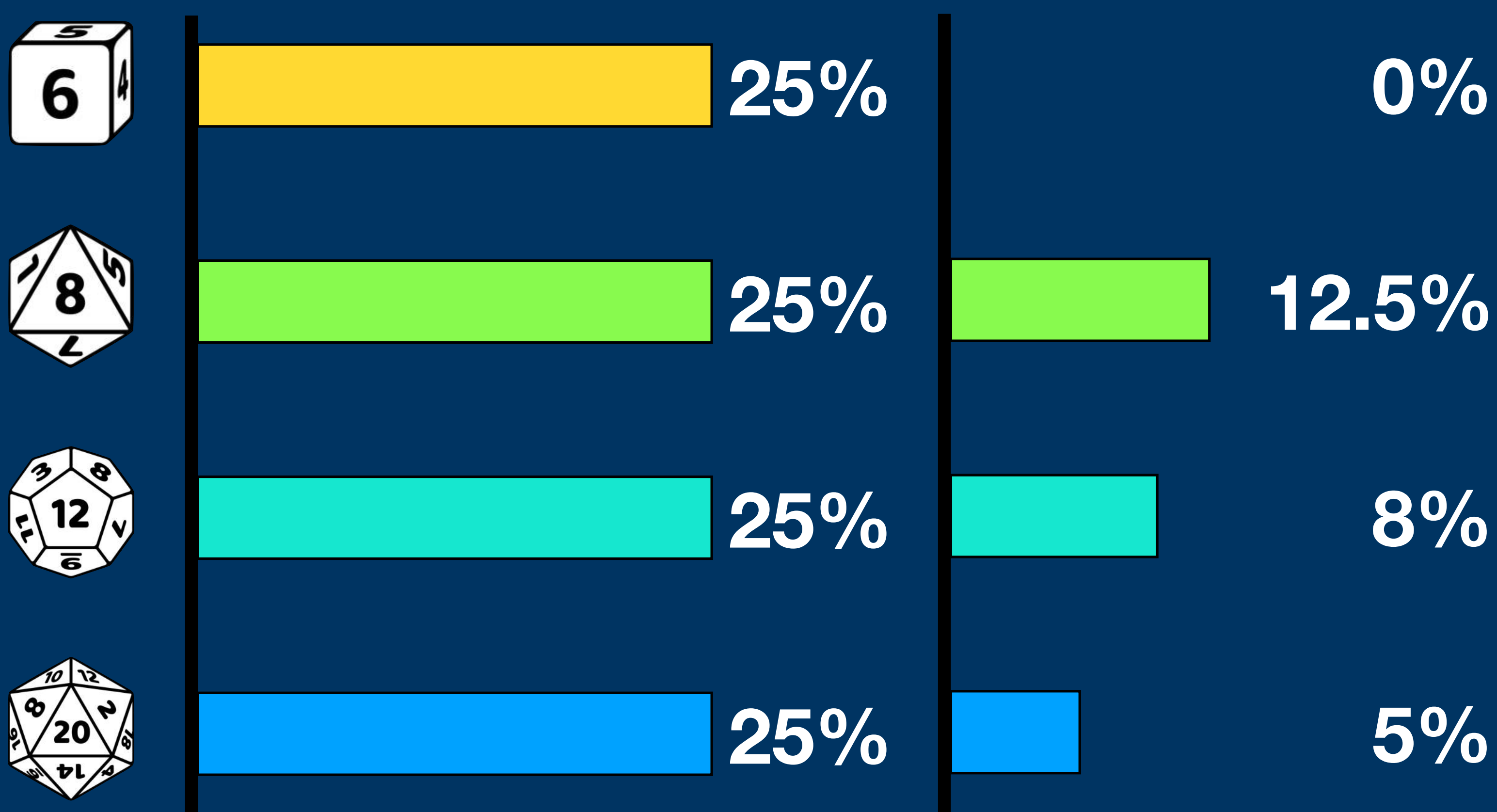
With the bayesian approach



Case study – Which dice did I choose ?

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

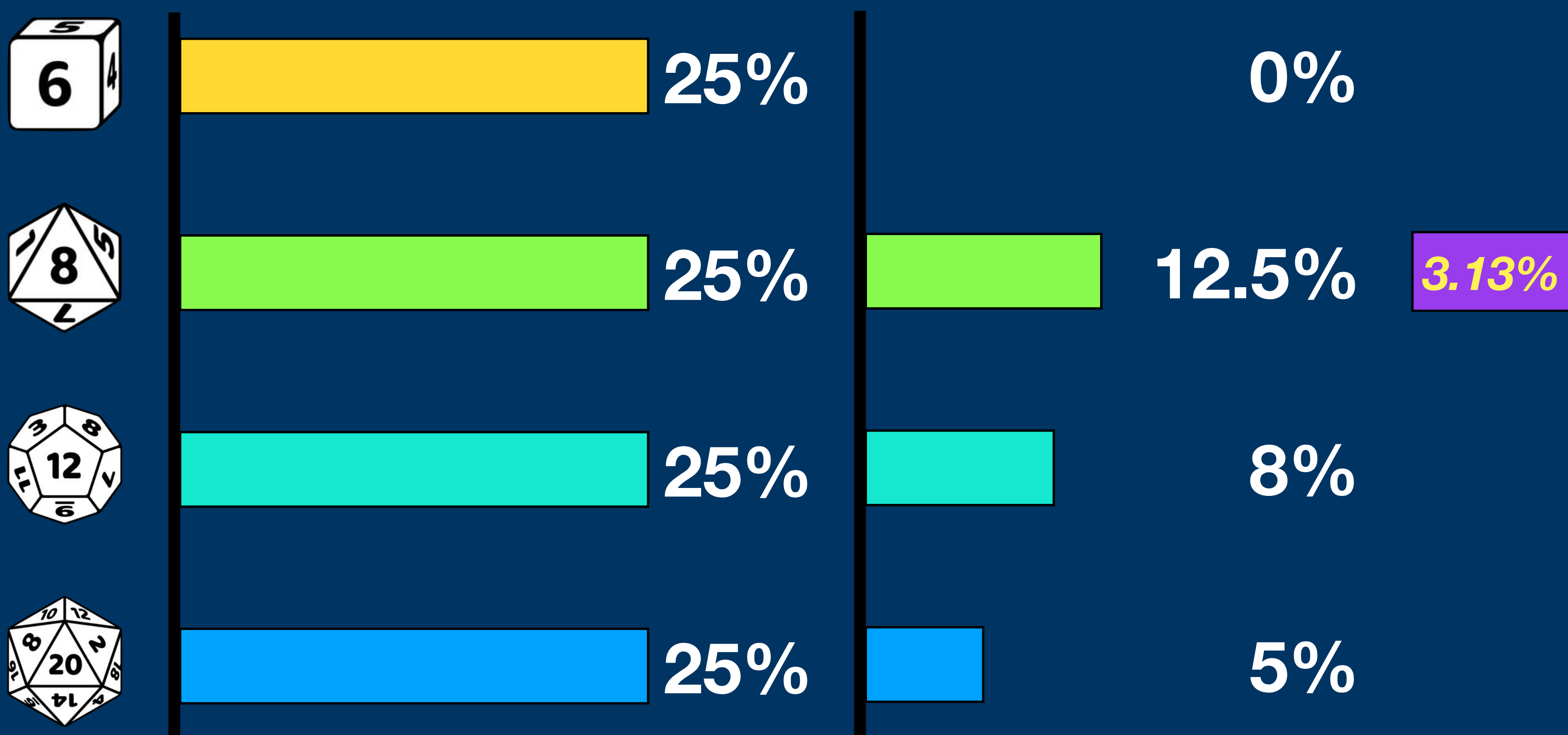
With the bayesian approach



Case study – Which dice did I choose ?

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

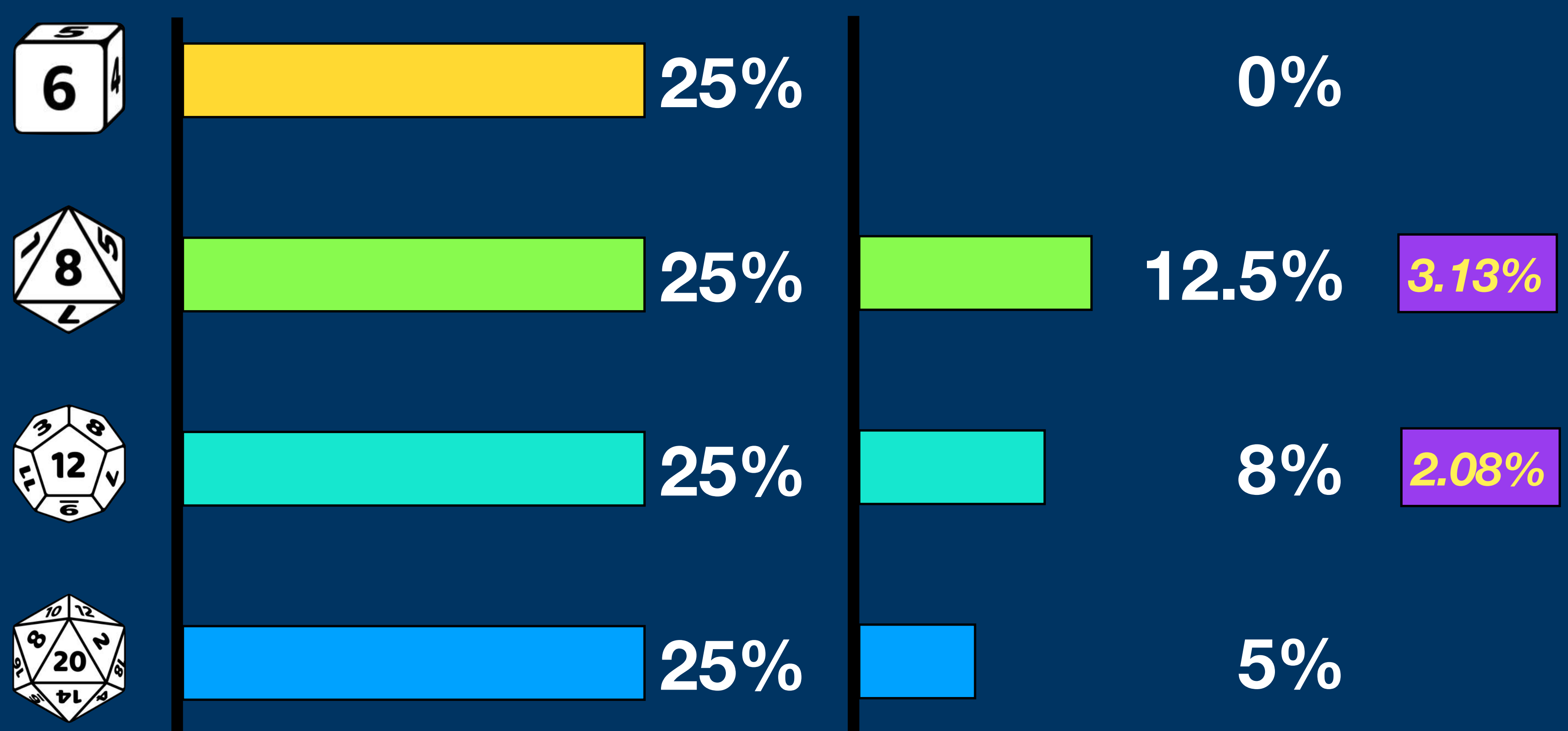
With the bayesian approach



Case study – Which dice did I choose ?

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

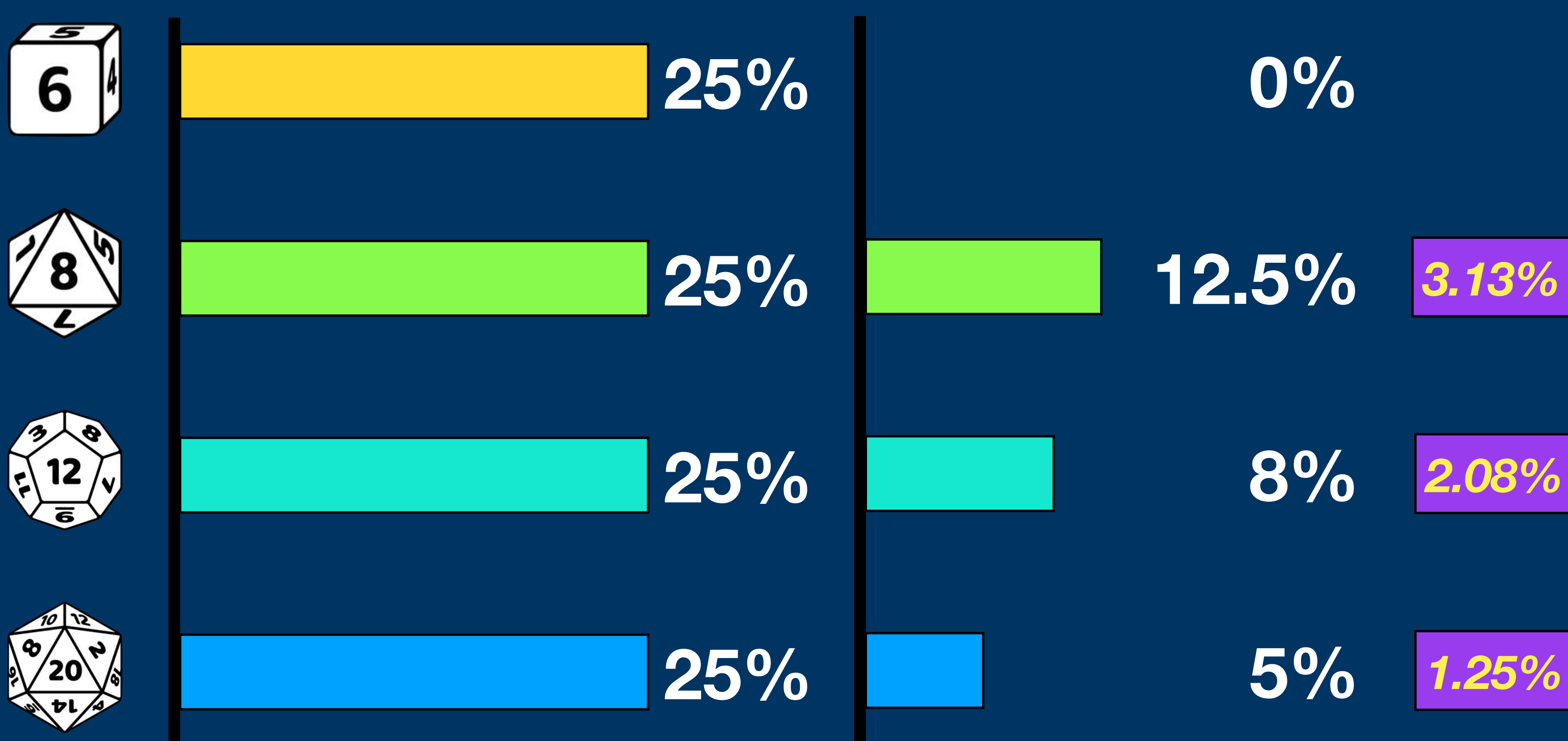
With the bayesian approach



Case study – Which dice did I choose ?

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

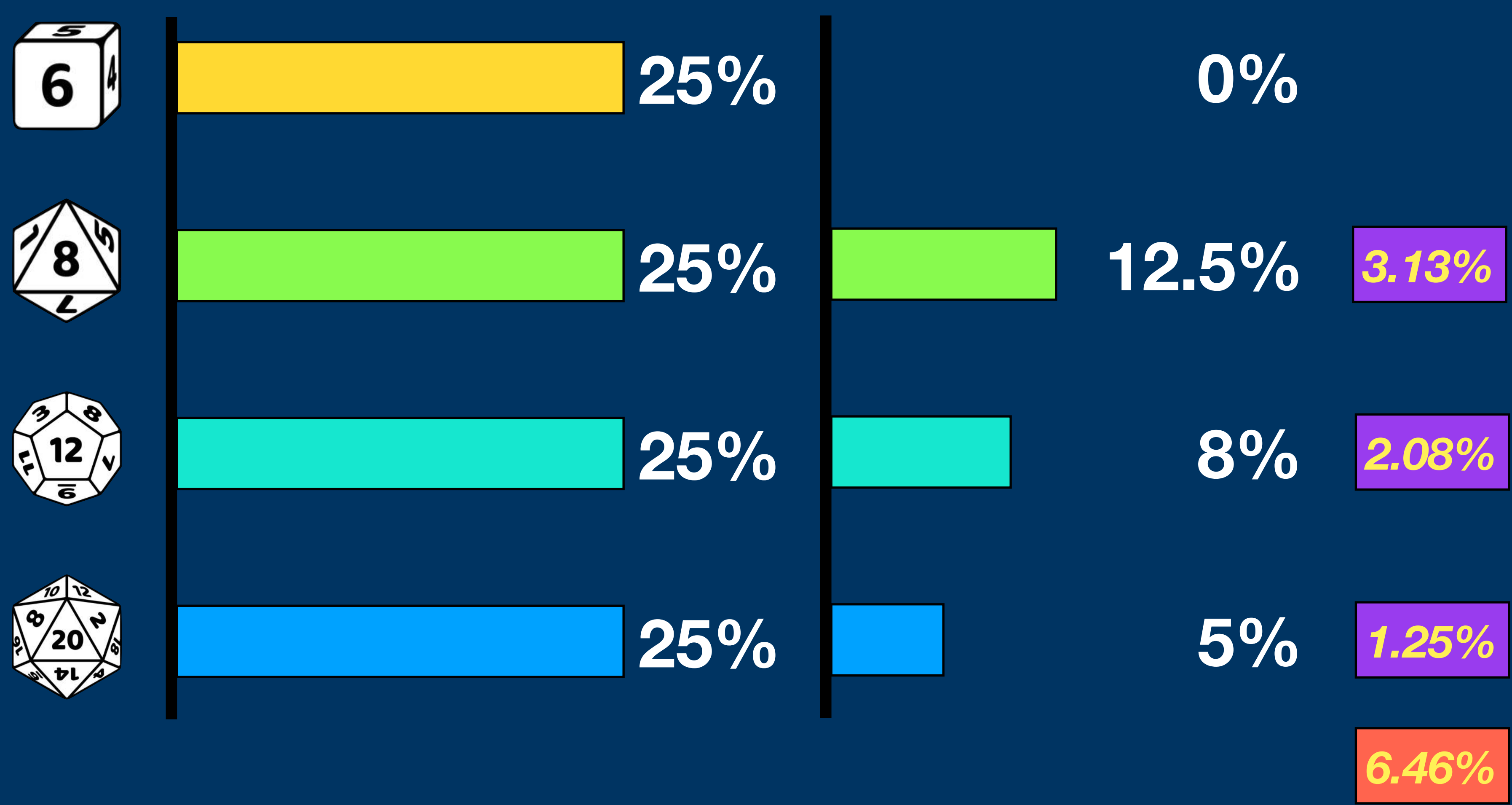
With the bayesian approach



Case study – Which dice did I choose ?

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

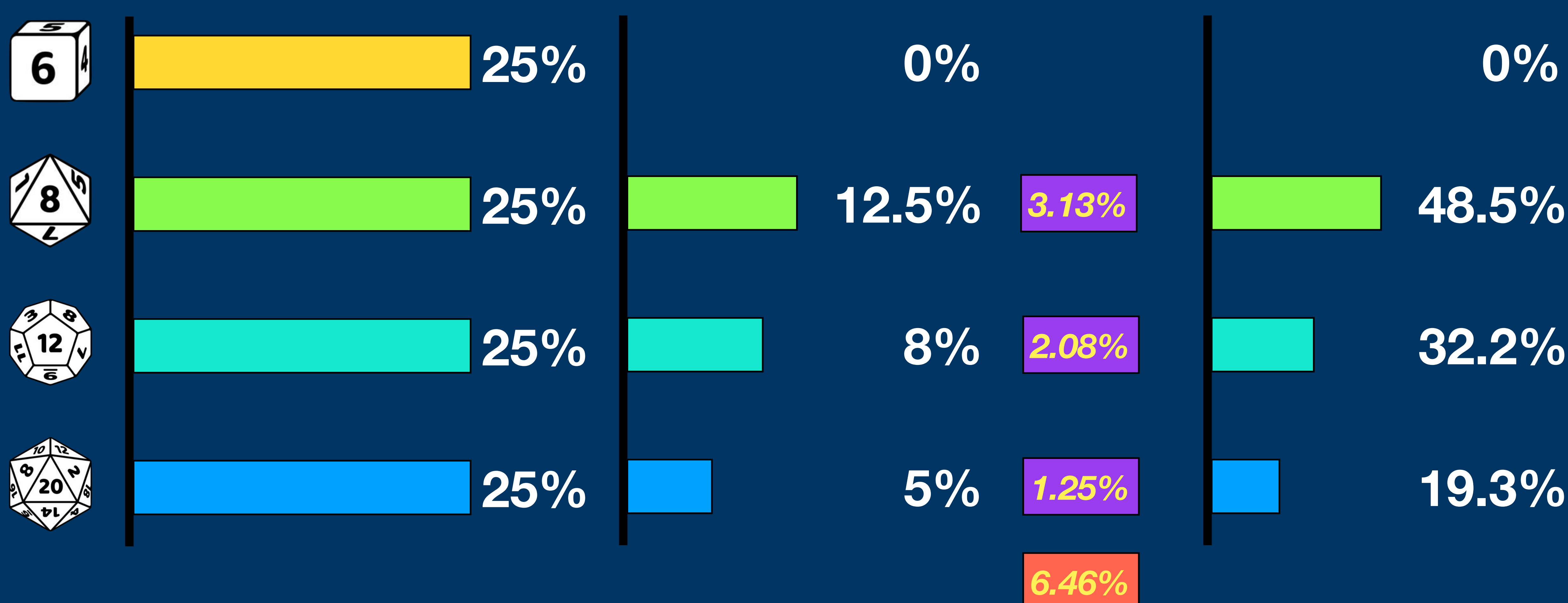
With the bayesian approach



Case study – Which dice did I choose ?

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

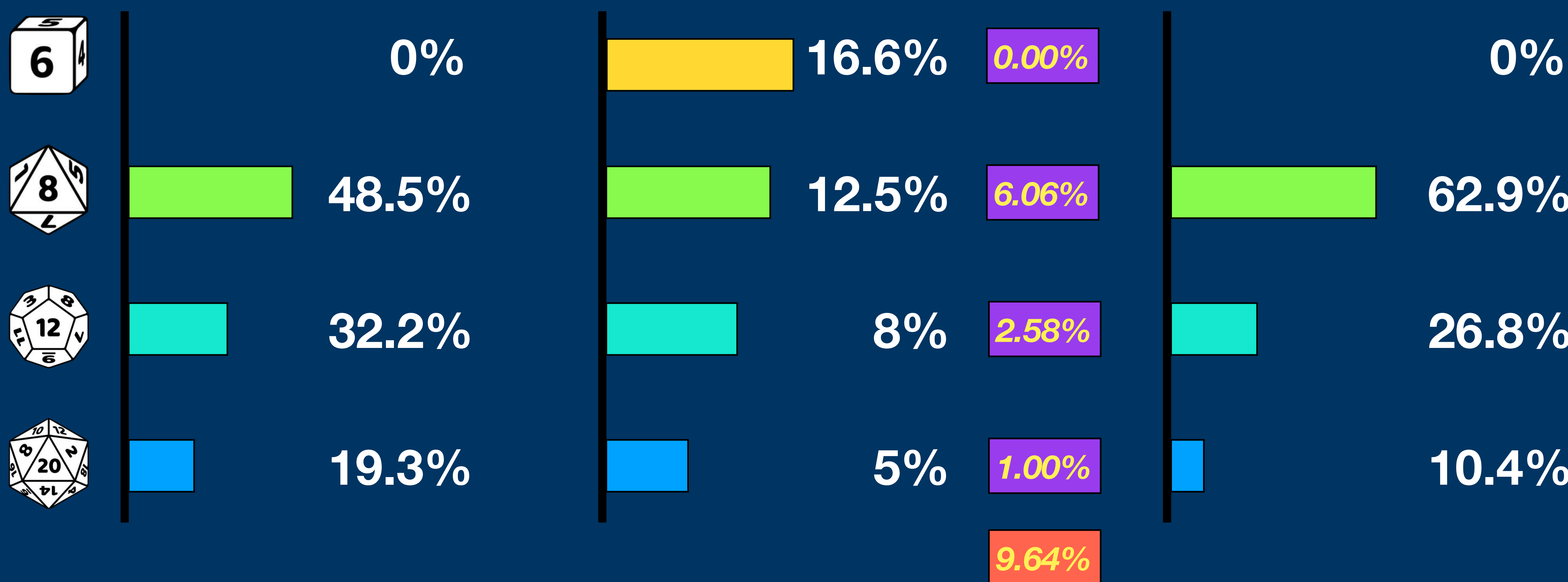
With the bayesian approach



Case study – Which dice did I choose ?

then I draw 2

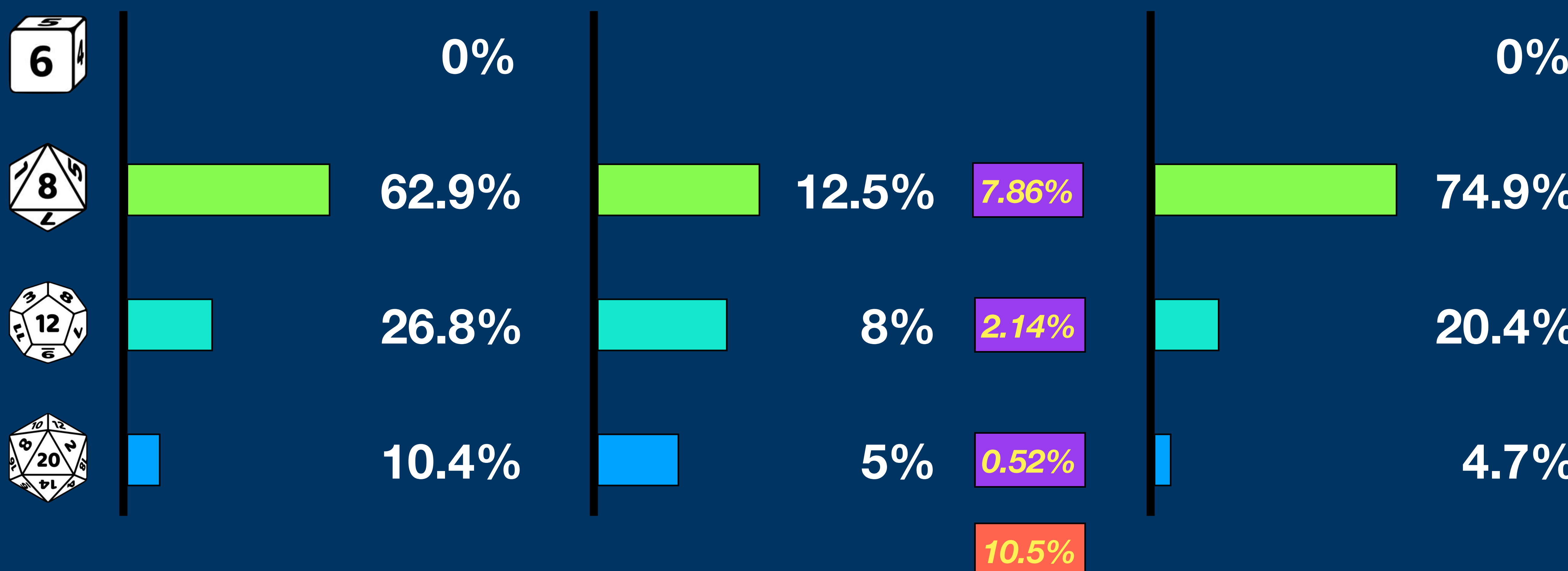
With the bayesian approach



Case study – Which dice did I choose ?

then I draw 8

With the bayesian approach

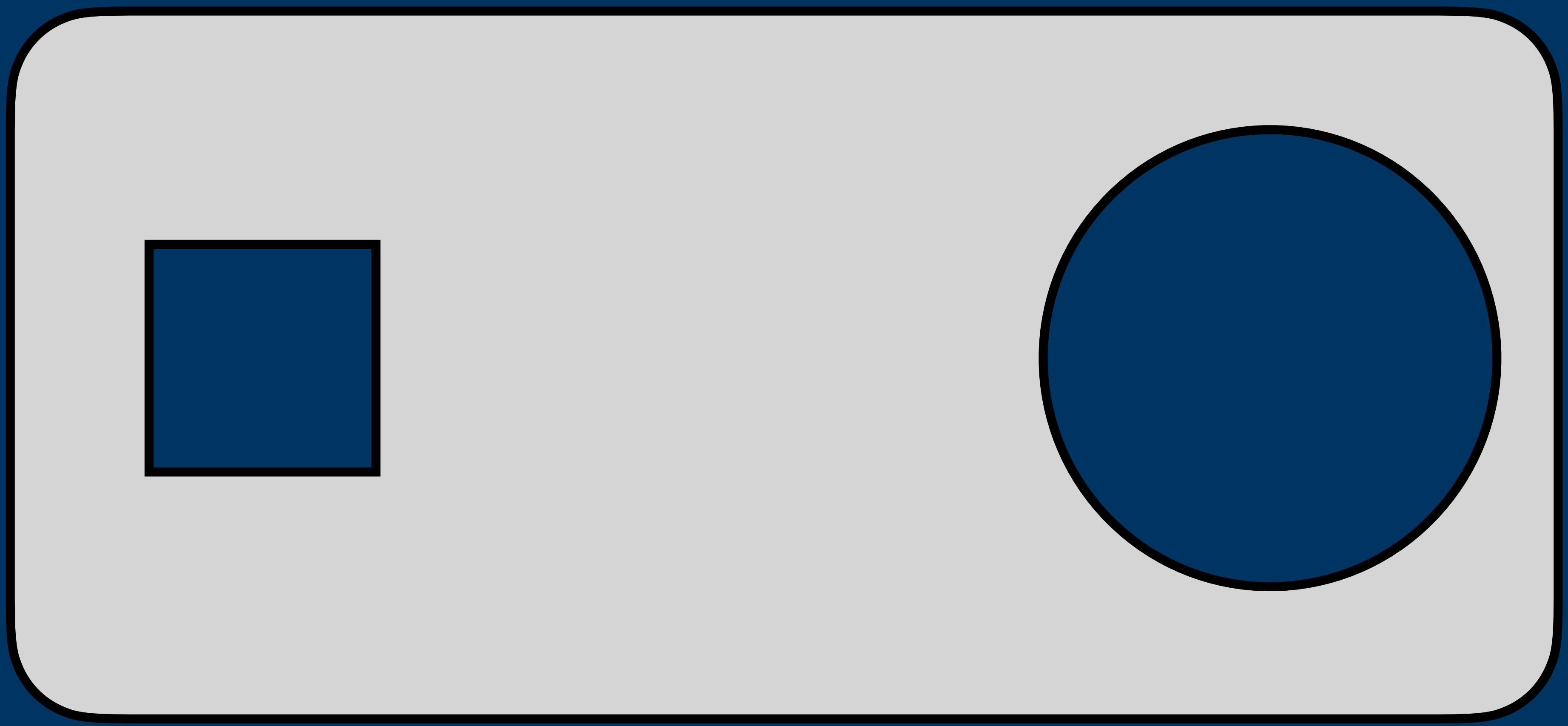


Part 1 – Prerequisites

Monte Carlo Markov Chains (MCMC)

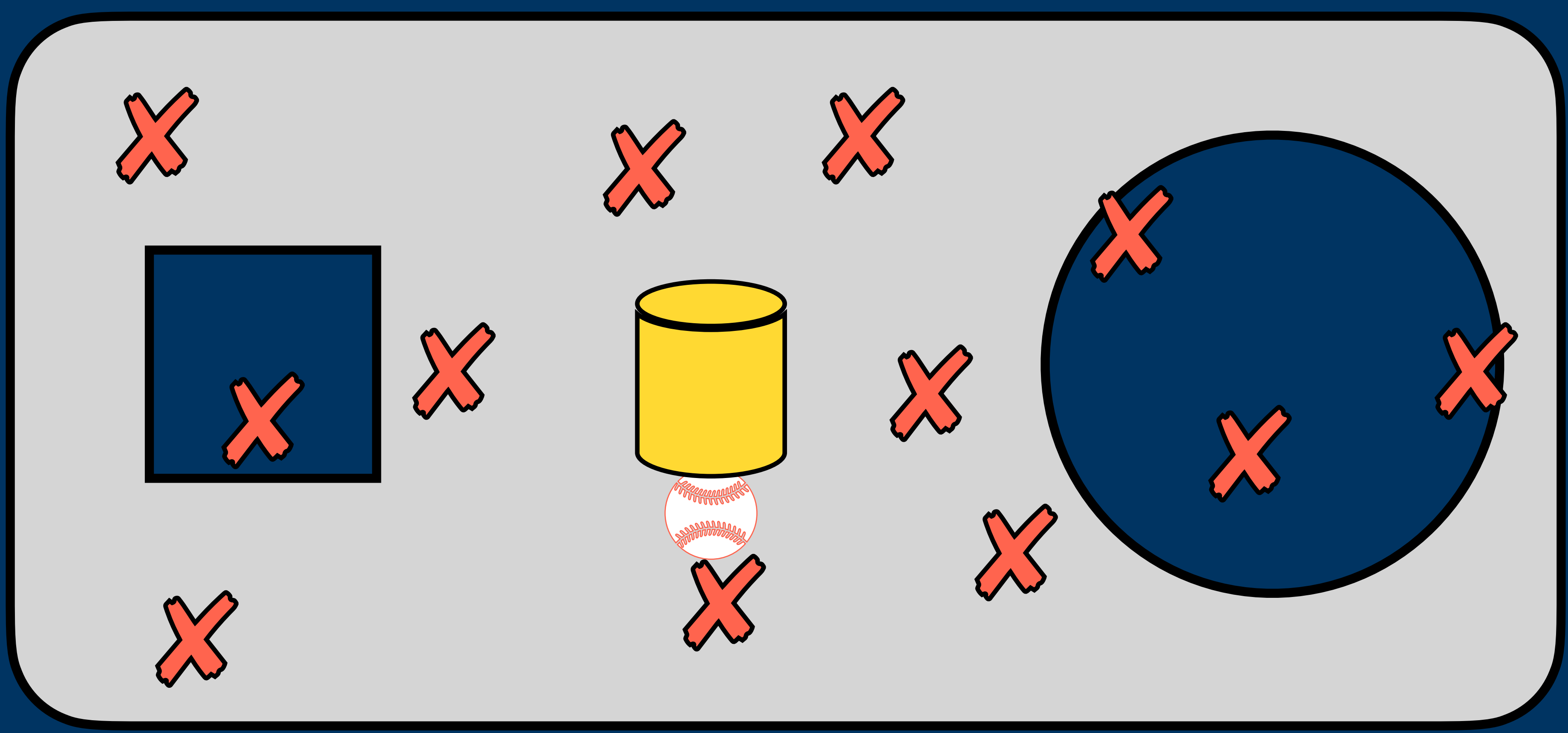
MCMC is synonymous with randomness. It involves simulations that evolve randomly. The term originates from Monaco, known for its casinos and gambling.

Consider the following scenario: A table with two holes — a round one and a square one — where we randomly throw a ball several times.

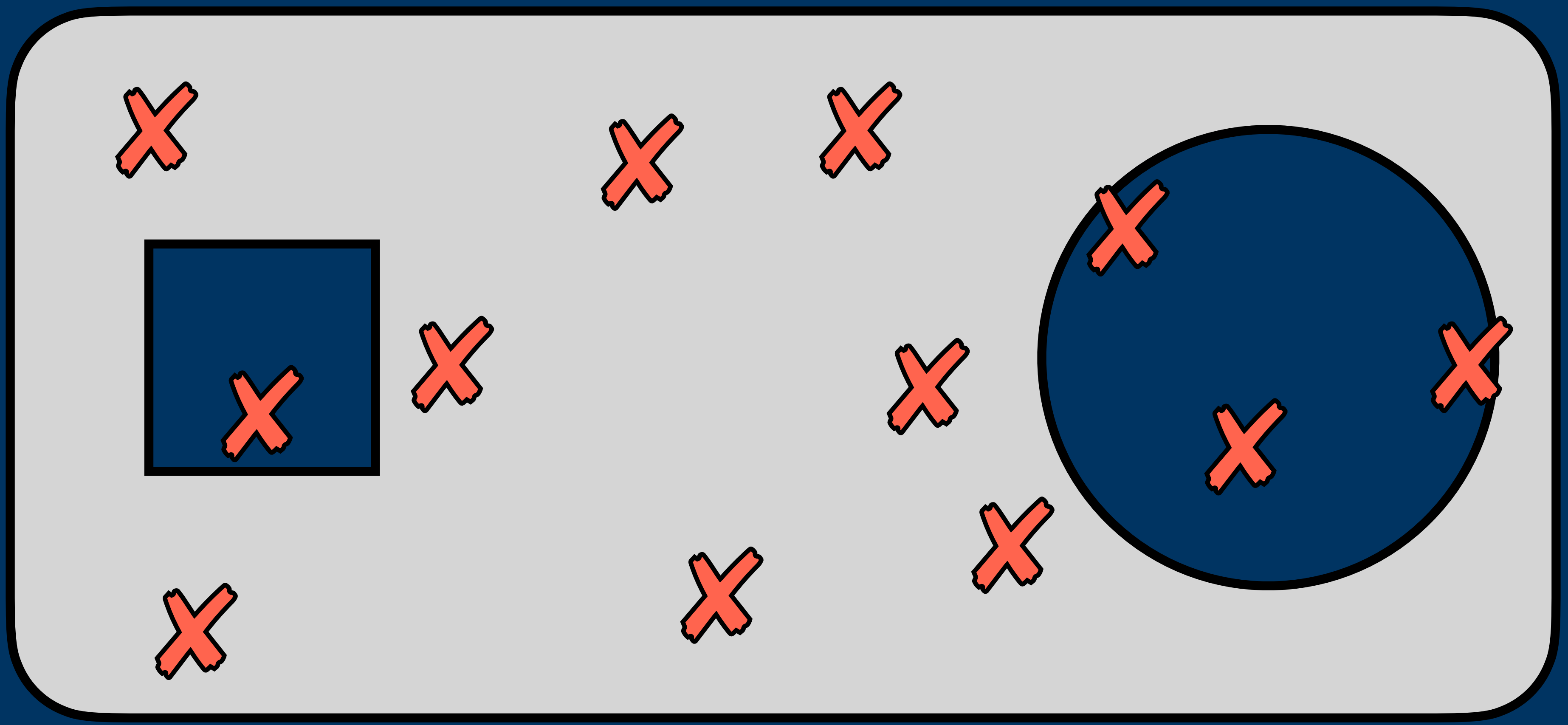


MCMC is synonymous with randomness. It involves simulations that evolve randomly. The term originates from Monaco, known for its casinos and gambling.

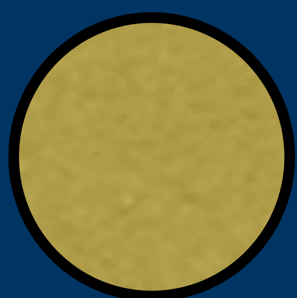
Consider the following scenario: A table with two holes — a round one and a square one — where we randomly throw a ball several times.



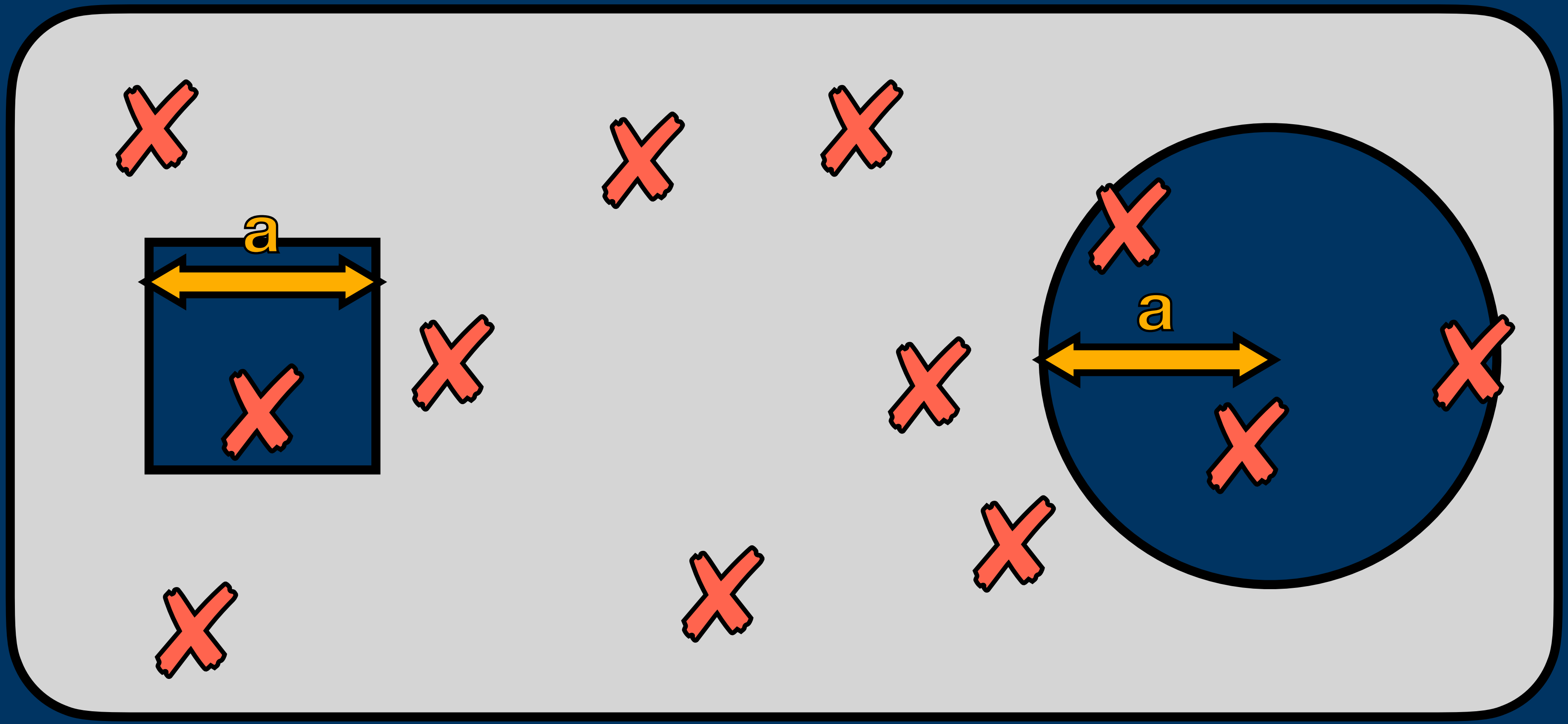
If we calculate the ratio of the number of balls inside each geometric shape, we will find:  $= \pi$



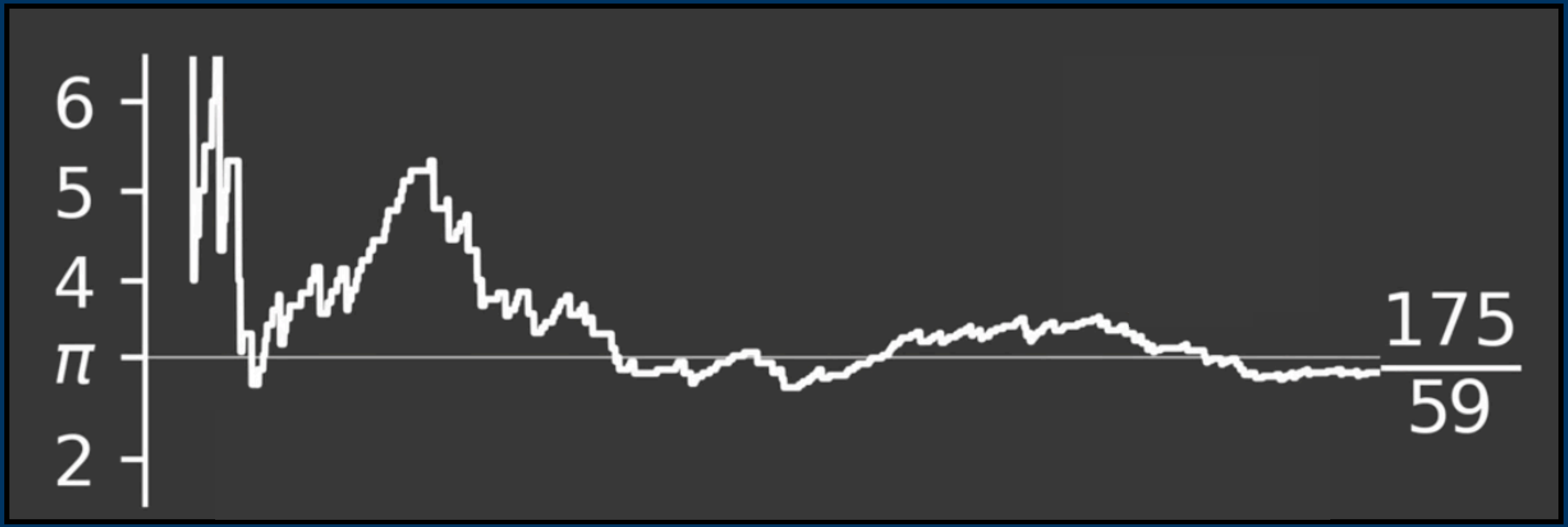
If we calculate the ratio of the number of balls inside each geometric shape, we will find:  $= \pi$

 $= \pi \cdot a^2$

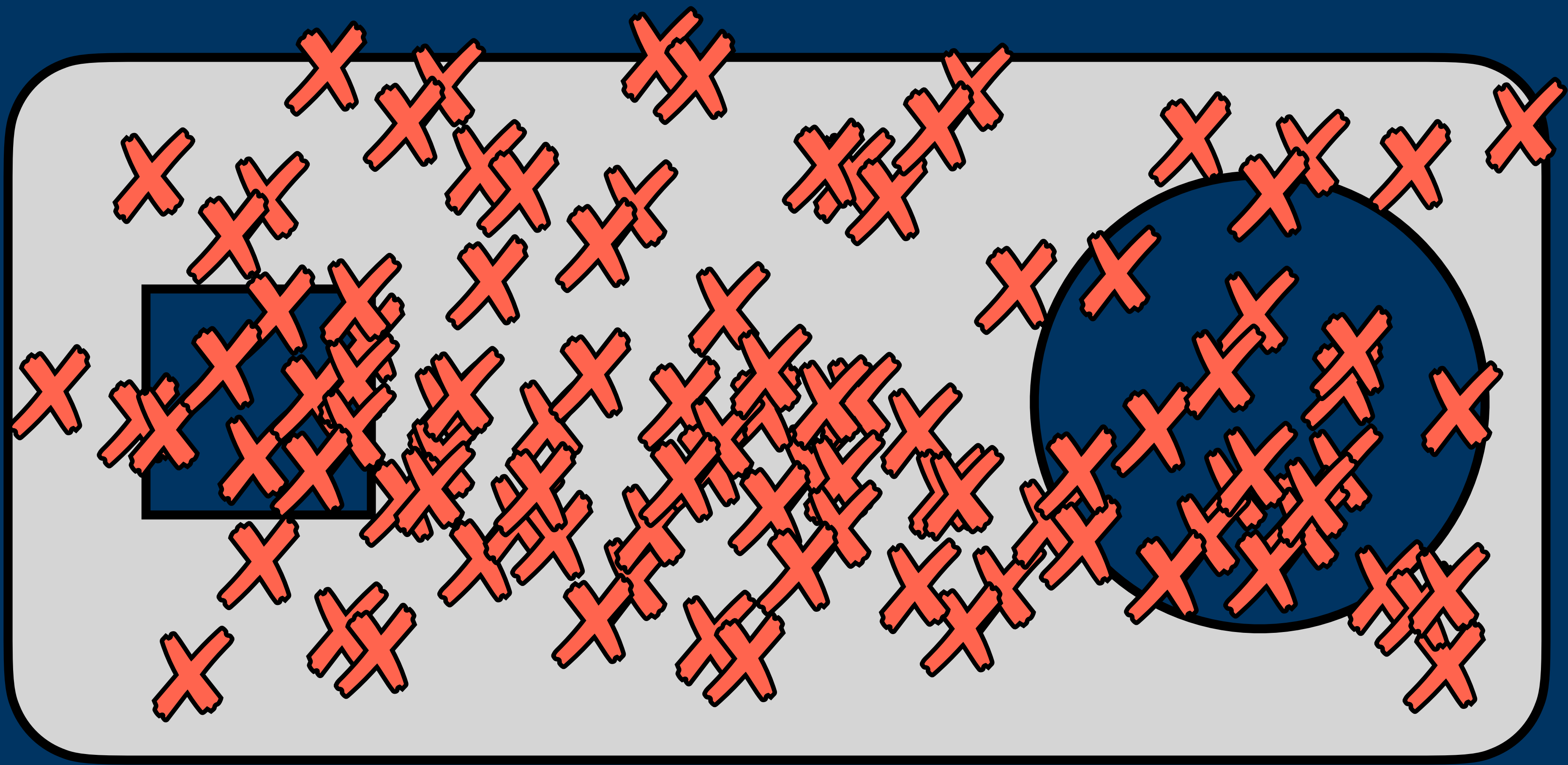
 $= a^2$



If we repeat this experiment a huge number of times, we can make a trace plot of our results:



Law of large numbers



Part 2 – Applications in Biology

Case study – impact of alcohol on heart

A data.frame with 21 observations on 5 variables:

country:
alcohol: liters alcohol from wine, per capita
deaths: deaths per 100,000
heart: heart disease dths per 100,000
liver: liver disease dths per 100,000

country	alcohol	deaths	heart	liver
Australia	2,5	785	211	15,3
Austria	3,9	863	167	45,6
Belg/Lux	2,9	883	131	20,7
Canada	2,4	793	191	16,4
Denmark	2,9	971	220	23,9
Finland	0,8	970	297	19,0
France	9,1	751	71	37,9
Iceland	0,8	743	211	11,2
Ireland	0,7	1000	300	6,5
Israel	0,6	834	183	13,7
Italy	7,9	775	107	42,2
Japan	1,5	680	36	23,2
Netherlands	1,8	773	167	9,2
New Zealand	1,9	916	266	7,7
Norway	0,8	806	227	12,2
Spain	6,5	724	86	36,4
Sweden	1,6	743	207	11,2
Switzerland	5,8	693	115	20,3
UK	1,3	941	285	10,3
US	1,2	926	199	22,1
West Germany	2,7	861	172	36,7

Is there a risk if I drink too much wine ?

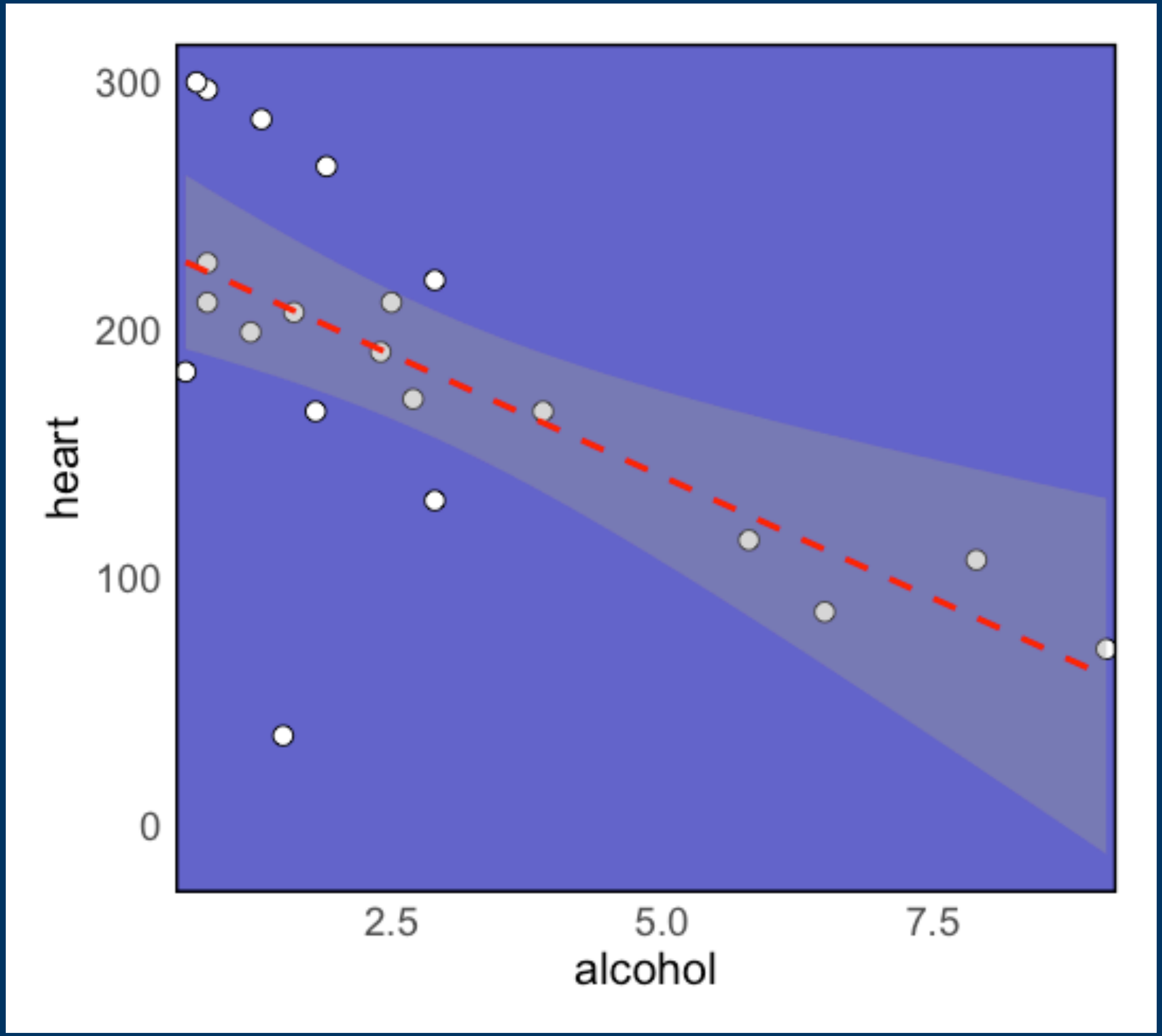
Case study – impact of alcohol on heart

A data.frame with 21 observations on 5 variables:

country:
alcohol: liters alcohol from wine, per capita
deaths: deaths per 100,000
heart: heart disease dths per 100,000
liver: liver disease dths per 100,000

Is there a risk if I drink too much wine ?

country	alcohol	deaths	heart	liver
Australia	2,5	785	211	15,3
Austria	3,9	863	167	45,6
Belg/Lux	2,9	883	131	20,7
Canada	2,4	793	191	16,4
Denmark	2,9	971	220	23,9
Finland	0,8	970	297	19,0
France	9,1	751	71	37,9
Iceland	0,8	743	211	11,2
Ireland	0,7	1000	300	6,5
Israel	0,6	834	183	13,7
Italy	7,9	775	107	42,2
Japan	1,5	680	36	23,2
Netherlands	1,8	773	167	9,2
New Zealand	1,9	916	266	7,7
Norway	0,8	806	227	12,2
Spain	6,5	724	86	36,4
Sweden	1,6	743	207	11,2
Switzerland	5,8	693	115	20,3
UK	1,3	941	285	10,3
US	1,2	926	199	22,1
West Germany	2,7	861	172	36,7



Case study – impact of alcohol on heart

With the frequentist approach

```
> summary(lm(heart ~ alcohol, data = wooldridge::wine))

Call:
lm(formula = heart ~ alcohol, data = wooldridge::wine)

Residuals:
     Min       1Q   Median       3Q      Max 
-173.623  -16.528   -0.655   23.346   74.631 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  239.147     19.032  12.565 1.19e-10 ***
alcohol      -19.683      5.121  -3.843   0.0011 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56.3 on 19 degrees of freedom
Multiple R-squared:  0.4374,	Adjusted R-squared:  0.4078 
F-statistic: 14.77 on 1 and 19 DF,  p-value: 0.001096
```

Easy code (only one line)

Case study – impact of alcohol on heart

With the frequentist approach

```
> summary(lm(heart ~ alcohol, data = wooldridge::wine))

Call:
lm(formula = heart ~ alcohol, data = wooldridge::wine)

Residuals:
     Min       1Q   Median       3Q      Max 
-173.623  -16.528   -0.655   23.346   74.631 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  239.147     19.032  12.565 1.19e-10 ***
alcohol      -19.683      5.121  -3.843   0.0011 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 56.3 on 19 degrees of freedom
Multiple R-squared:  0.4374,	Adjusted R-squared:  0.4078 
F-statistic: 14.77 on 1 and 19 DF,  p-value: 0.001096
```

Easy code (only one line)

My estimates with the corresponding errors

Case study – impact of alcohol on heart

With the frequentist approach

```
> summary(lm(heart ~ alcohol, data = wooldridge::wine))

Call:
lm(formula = heart ~ alcohol, data = wooldridge::wine)

Residuals:
     Min       1Q   Median       3Q      Max
-173.623  -16.528   -0.655   23.346   74.631

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  239.147     19.032  12.565 1.19e-10 ***
alcohol      -19.683      5.121  -3.843   0.0011 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56.3 on 19 degrees of freedom
Multiple R-squared:  0.4374,	Adjusted R-squared:  0.4078
F-statistic: 14.77 on 1 and 19 DF,  p-value: 0.001096
```

Easy code (only one line)

My estimates with the corresponding errors

My R^2

Case study – impact of alcohol on heart

With the frequentist approach

```
> summary(lm(heart ~ alcohol, data = wooldridge::wine))

Call:
lm(formula = heart ~ alcohol, data = wooldridge::wine)

Residuals:
     Min       1Q   Median       3Q      Max 
-173.623  -16.528   -0.655   23.346   74.631 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  239.147     19.032  12.565 1.19e-10 ***
alcohol      -19.683      5.121  -3.843   0.0011 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56.3 on 19 degrees of freedom
Multiple R-squared:  0.4374,	Adjusted R-squared:  0.4078 
F-statistic: 14.77 on 1 and 19 DF,  p-value: 0.001096
```

Easy code (only one line)

My estimates with the corresponding errors

My R^2

Even my p-value

Case study – impact of alcohol on heart

With the frequentist approach

As stated before, It's easy to compute. We quickly conclude (~~while omitting to mention the limited sample size of only 21 observations~~) that French 🇫🇷, followed by Italians 🇮🇹 and Spanish 🇪🇸 can continue to drink without any risk and we submit our findings for publication in *Nature*.

Case study – impact of alcohol on heart

With the bayesian approach

As stated before, It's easy to compute. We quickly conclude (~~while omitting to mention the limited sample size of only 21 observations~~) that French 🇫🇷, followed by Italians 🇮🇹 and Spanish 🇪🇸 can continue to drink without any risk and we submit our findings for publication in *Nature*.

But you've been rejected because you did not consider a major previous study…

Case study – impact of alcohol on heart

With the bayesian approach

Before proceeding further, there are two types of "teams" among scientists. There are those who utilize Bayesian methods to enhance the accuracy of calculations, ensuring robustness while maintaining complete objectivity (*i.e.,* non-informative priors). Then, there are those who incorporate prior information to refine posterior outputs.

In our case, we align with the latter group of scientists. We have been advised by the editor about a study stating that the slope found in a previous study was more likely -15 with a standard error of 5.

Case study – impact of alcohol on heart

With the bayesian approach

```r
reg  <- brms::brm(formula = heart ~ alcohol, data = wooldridge::wine,
                  chains = 3, iter = 20000, warmup = 2500, cores = 3,
                  prior = prior(normal(-15, 5), class = "b", coef = "alcohol"))
```

Case study – impact of alcohol on heart

With the bayesian approach

Exact same line of code as previously!

```
reg  ← brms::brm(formula = heart ~ alcohol, data = wooldridge::wine,
                 chains = 3, iter = 20000, warmup = 2500, cores = 3,
                 prior = prior(normal(-15, 5), class = "b", coef = "alcohol"))
```

Case study – impact of alcohol on heart

With the bayesian approach

Exact same line of code as previously!

The number of MCMC we want ***(2 to 5 recommended)***

```
reg  ← brms::brm(formula = heart ~ alcohol, data = wooldridge::wine,
                 chains = 3, iter = 20000, warmup = 2500, cores = 3,
                 prior = prior(normal(-15, 5), class = "b", coef = "alcohol"))
```

Case study – impact of alcohol on heart

With the bayesian approach

Exact same line of code as previously!

The number of MCMC we want ***(2 to 5 recommended)***

The number of calculs you want by chain

```r
reg  <- brms::brm(formula = heart ~ alcohol, data = wooldridge::wine,
                  chains = 3, iter = 20000, warmup = 2500, cores = 3,
                  prior = prior(normal(-15, 5), class = "b", coef = "alcohol"))
```

Case study – impact of alcohol on heart

With the bayesian approach

Exact same line of code as previously!

The number of MCMC we want *(2 to 5 recommended)*

The number of calculs you want by chain

```
reg ← brms::brm(formula = heart ~ alcohol, data = wooldridge::wine,
                chains = 3, iter = 20000, warmup = 2500, cores = 3,
                prior = prior(normal(-15, 5), class = "b", coef = "alcohol"))
```

The number of calculs you want
to discard at the beginning

Case study – impact of alcohol on heart

With the bayesian approach

Exact same line of code as previously!

The number of MCMC we want *(2 to 5 recommended)*

The number of calculs you want by chain

```
reg ← brms::brm(formula = heart ~ alcohol, data = wooldridge::wine,
                chains = 3, iter = 20000, warmup = 2500, cores = 3,
                prior = prior(normal(-15, 5), class = "b", coef = "alcohol"))
```

The number of calculs you want to discard at the beginning

The prior formulation for the slope b from the variable “alcohol”

Case study – impact of alcohol on heart

With the bayesian approach

```
> reg
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: heart ~ alcohol
   Data: wooldridge::wine (Number of observations: 21)
  Draws: 3 chains, each with iter = 20000; warmup = 2500; thin = 1;
         total post-warmup draws = 52500

Population-Level Effects:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   232.92     16.04   200.74   264.09 1.00    46557    34962
alcohol     -17.29      3.62   -24.25   -10.02 1.00    44605    36800

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    57.73      9.45    42.66    79.39 1.00    41533    35356

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
> brms::bayes_R2(reg)
    Estimate Est.Error      Q2.5     Q97.5
R2 0.3638462 0.1003611 0.1450943 0.5310754
```

Case study – impact of alcohol on heart

With the bayesian approach

```
> reg
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: heart ~ alcohol
   Data: wooldridge::wine (Number of observations: 21)
  Draws: 3 chains, each with iter = 20000; warmup = 2500; thin = 1;
         total post-warmup draws = 52500

Population-Level Effects:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   232.92     16.04   200.74   264.09 1.00    46557    34962
alcohol     -17.29      3.62   -24.25   -10.02 1.00    44605    36800

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    57.73      9.45    42.66    79.39 1.00    41533    35356

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

The total number of draws

```
> brms::bayes_R2(reg)
    Estimate Est.Error      Q2.5     Q97.5
R2 0.3638462 0.1003611 0.1450943 0.5310754
```

Case study – impact of alcohol on heart

With the bayesian approach

```
> reg
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: heart ~ alcohol
   Data: wooldridge::wine (Number of observations: 21)
  Draws: 3 chains, each with iter = 20000; warmup = 2500; thin = 1;
         total post-warmup draws = 52500

Population-Level Effects:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   232.92     16.04   200.74   264.09 1.00    46557    34962
alcohol     -17.29      3.62   -24.25   -10.02 1.00    44605    36800

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    57.73      9.45    42.66    79.39 1.00    41533    35356

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

The total number of draws

CI stands for credible interval and not confident interval

```
> brms::bayes_R2(reg)
    Estimate Est.Error      Q2.5     Q97.5
R2 0.3638462 0.1003611 0.1450943 0.5310754
```

Case study – impact of alcohol on heart

With the bayesian approach

```
> reg
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: heart ~ alcohol
   Data: wooldridge::wine (Number of observations: 21)
  Draws: 3 chains, each with iter = 20000; warmup = 2500; thin = 1;
         total post-warmup draws = 52500

Population-Level Effects:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   232.92     16.04   200.74   264.09 1.00    46557    34962
alcohol     -17.29      3.62   -24.25   -10.02 1.00    44605    36800

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    57.73      9.45    42.66    79.39 1.00    41533    35356

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
> brms::bayes_R2(reg)
    Estimate Est.Error      Q2.5     Q97.5
R2 0.3638462 0.1003611 0.1450943 0.5310754
```

The total number of draws

CI stands for credible interval and not confident interval

R_{hat} means it has converged

Case study – impact of alcohol on heart

With the bayesian approach

```
> reg
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: heart ~ alcohol
   Data: wooldridge::wine (Number of observations: 21)
  Draws: 3 chains, each with iter = 20000; warmup = 2500; thin = 1;
         total post-warmup draws = 52500

Population-Level Effects:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   232.92     16.04   200.74   264.09 1.00    46557    34962
alcohol     -17.29      3.62   -24.25   -10.02 1.00    44605    36800

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    57.73      9.45    42.66    79.39 1.00    41533    35356

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

The total number of draws

CI stands for credible interval and not confident interval

R_{hat} means it has converged

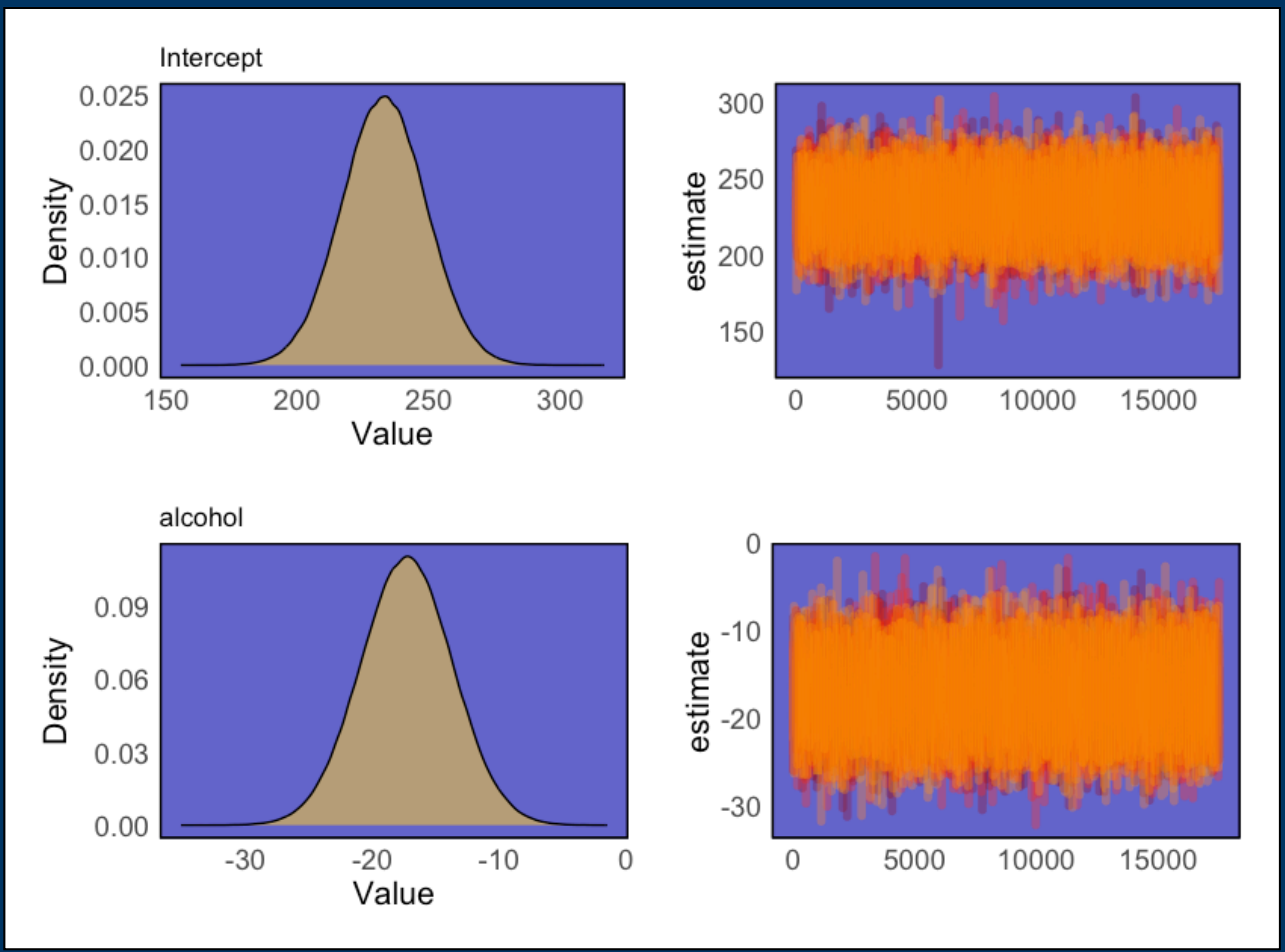
```
> brms::bayes_R2(reg)
    Estimate Est.Error      Q2.5     Q97.5
R2 0.3638462 0.1003611 0.1450943 0.5310754
```

Now you have a distrib. for your R^2

Case study – impact of alcohol on heart

With the bayesian approach

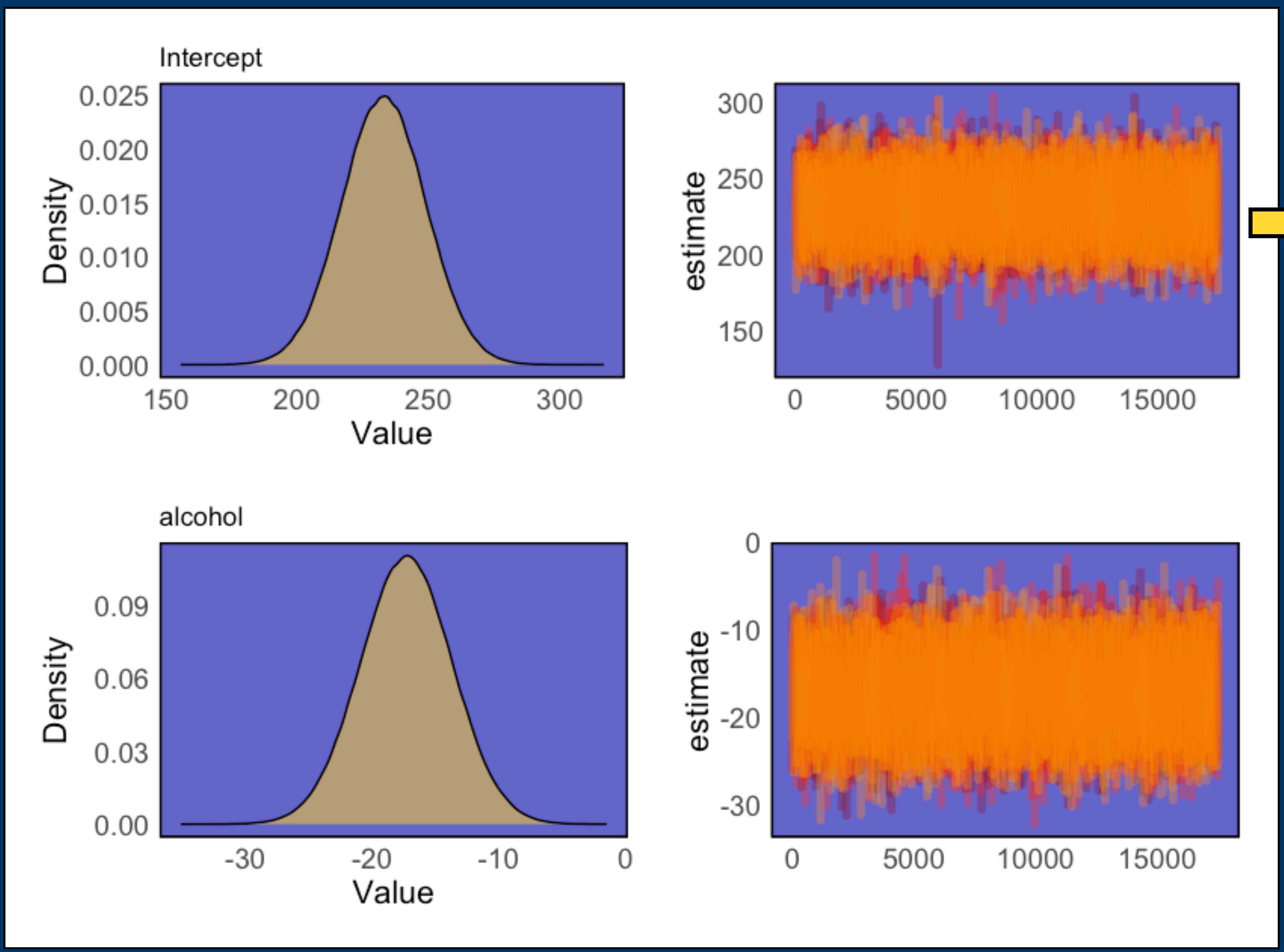
```
> plot(reg)
```



Case study – impact of alcohol on heart

With the bayesian approach

```
> plot(reg)
```

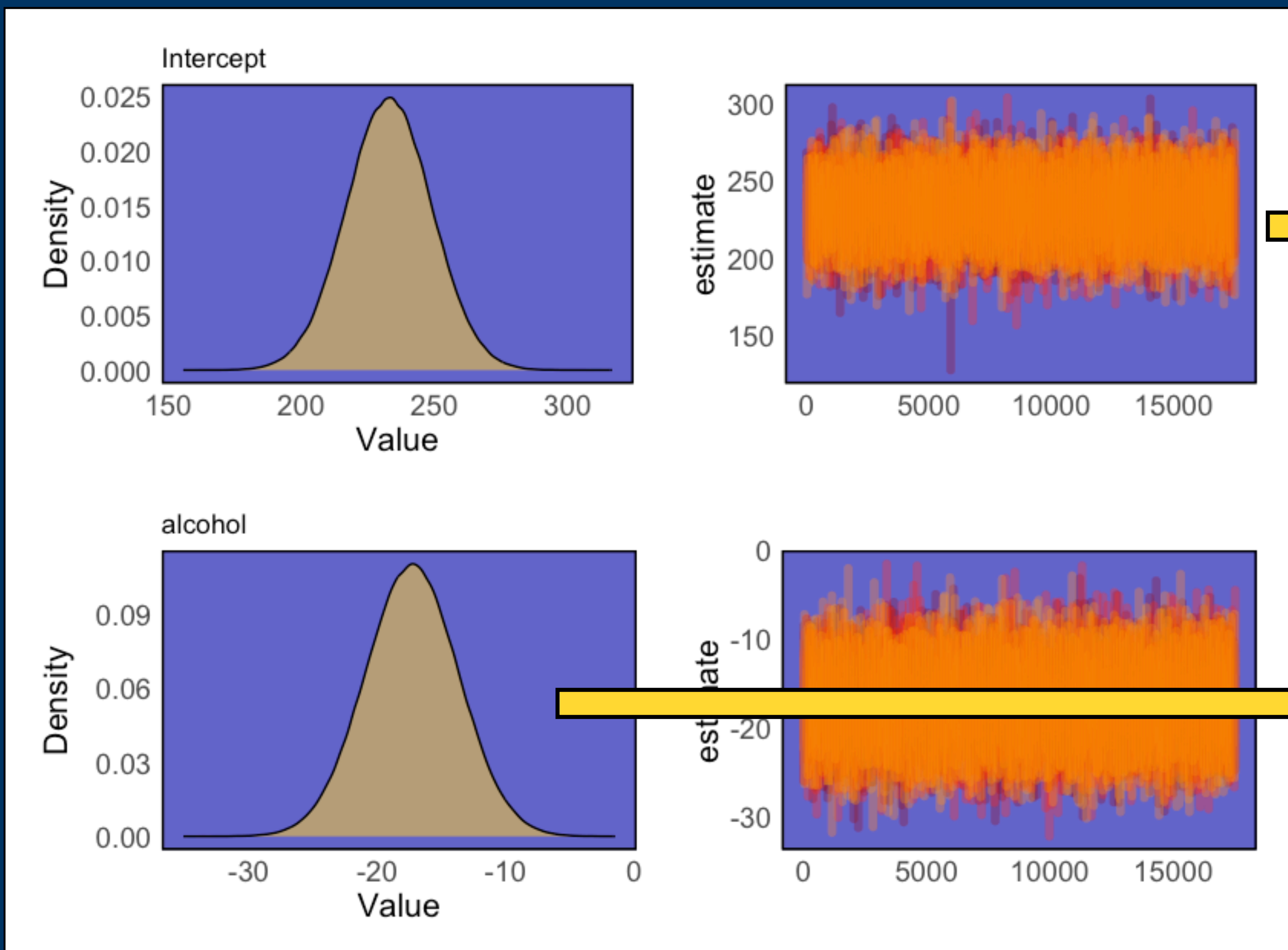


You can see the convergence of your MCMC chains

Case study – impact of alcohol on heart

With the bayesian approach

```
> plot(reg)
```

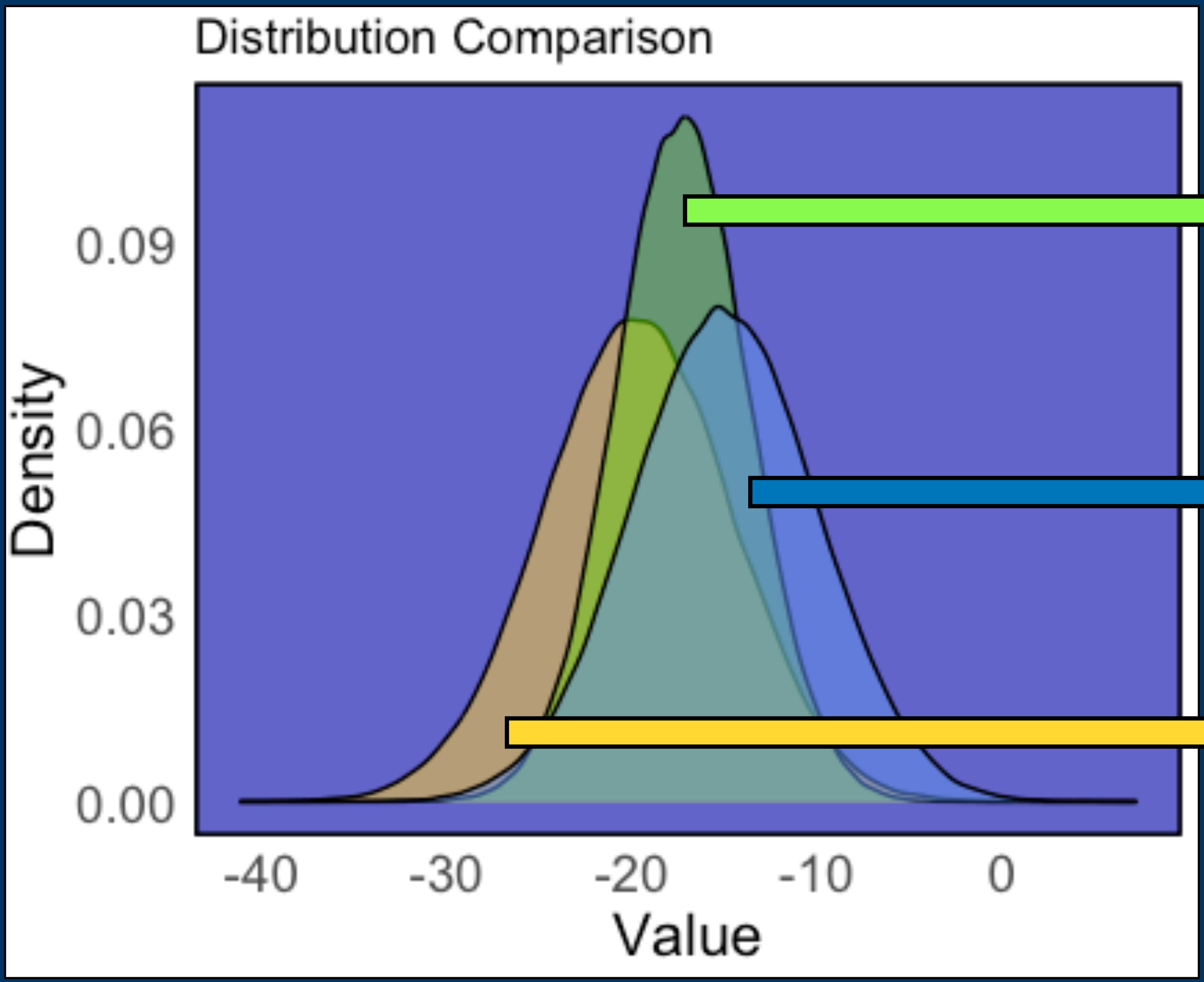


You can see the convergence of your MCMC chains

You can also see the distribution of your variables

Case study – impact of alcohol on heart

With the bayesian approach



Posterior N(-17.3, 3.6)

Prior N(-15, 5)

Likelihood N(-19.7, 5.1)

Case study – impact of alcohol on heart

With the bayesian approach

Now you're able to confirm that French 🇫🇷, followed by Italians 🇮🇹 and Spanish 🇪🇸, are safe and you can resubmit for publication in *Nature*. Kudos!

Case study – impact of alcohol on heart

With the bayesian approach

Now, imagine you want to argue that wine is harmful. Your goal is to completely shut down wine shops. You've realized that you can incorporate prior information that could alter the conclusion.

For instance, instead of using a prior distribution of N(-15, 5), you switch it to N(+15, 5).

Case study – impact of alcohol on heart

With the bayesian approach

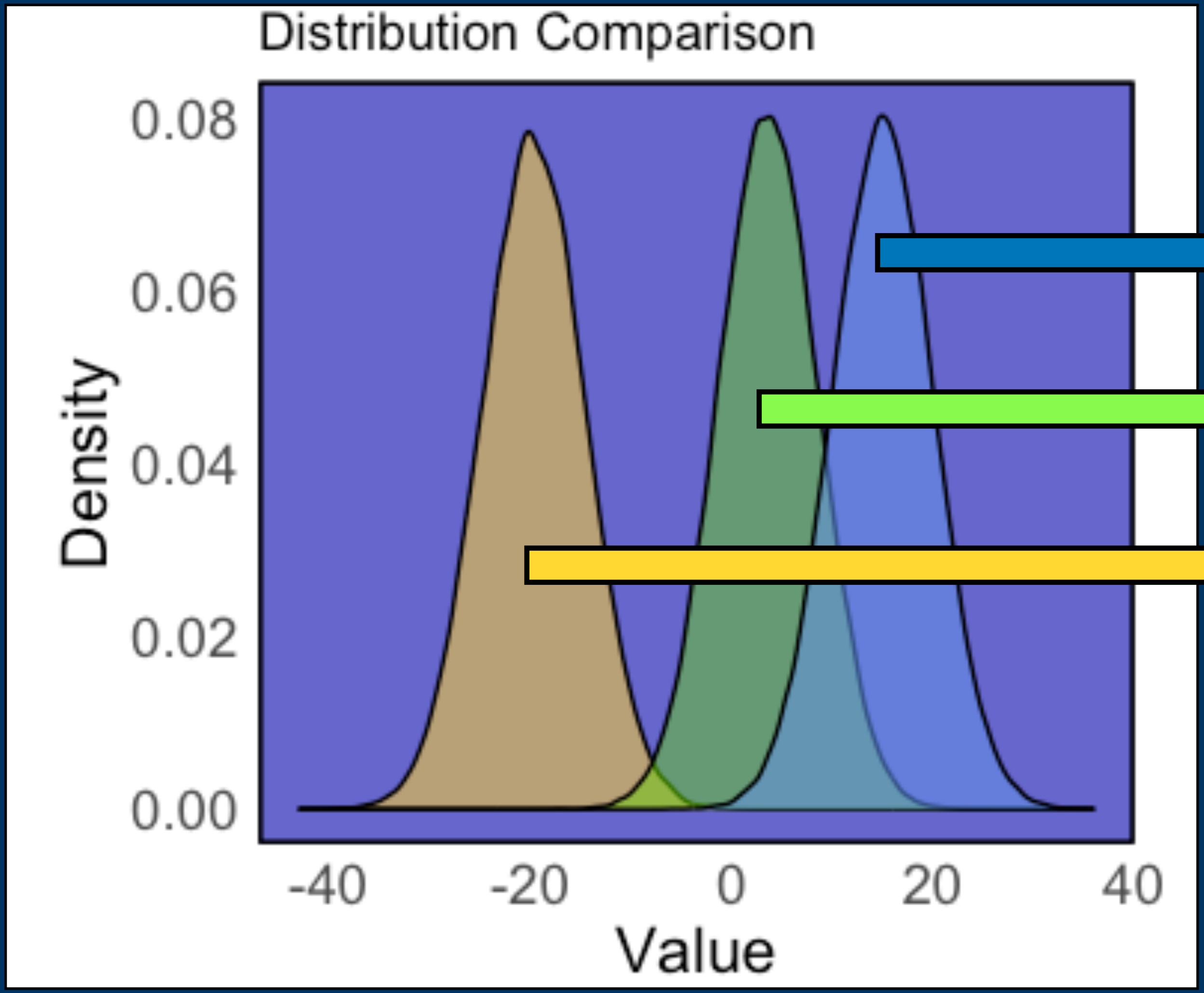
```
> reg
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: heart ~ alcohol
   Data: wooldridge::wine (Number of observations: 21)
  Draws: 3 chains, each with iter = 20000; warmup = 2500; thin = 1;
         total post-warmup draws = 52500

Population-Level Effects:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   174.00     21.66   128.93   213.89 1.00    33901    30686
alcohol       3.60      4.92    -5.72    13.46 1.00    29457    34014

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma    80.88     15.37    56.17   115.66 1.00    27885    33377

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
> brms::bayes_R2(reg)
     Estimate  Est.Error         Q2.5     Q97.5
R2 0.02777283 0.03003607 4.026603e-05 0.1057474
```

😱 Fortunately, your credible interval is both negative and positive

Explain only 2%

Case study – impact of alcohol on heart

With the bayesian approach



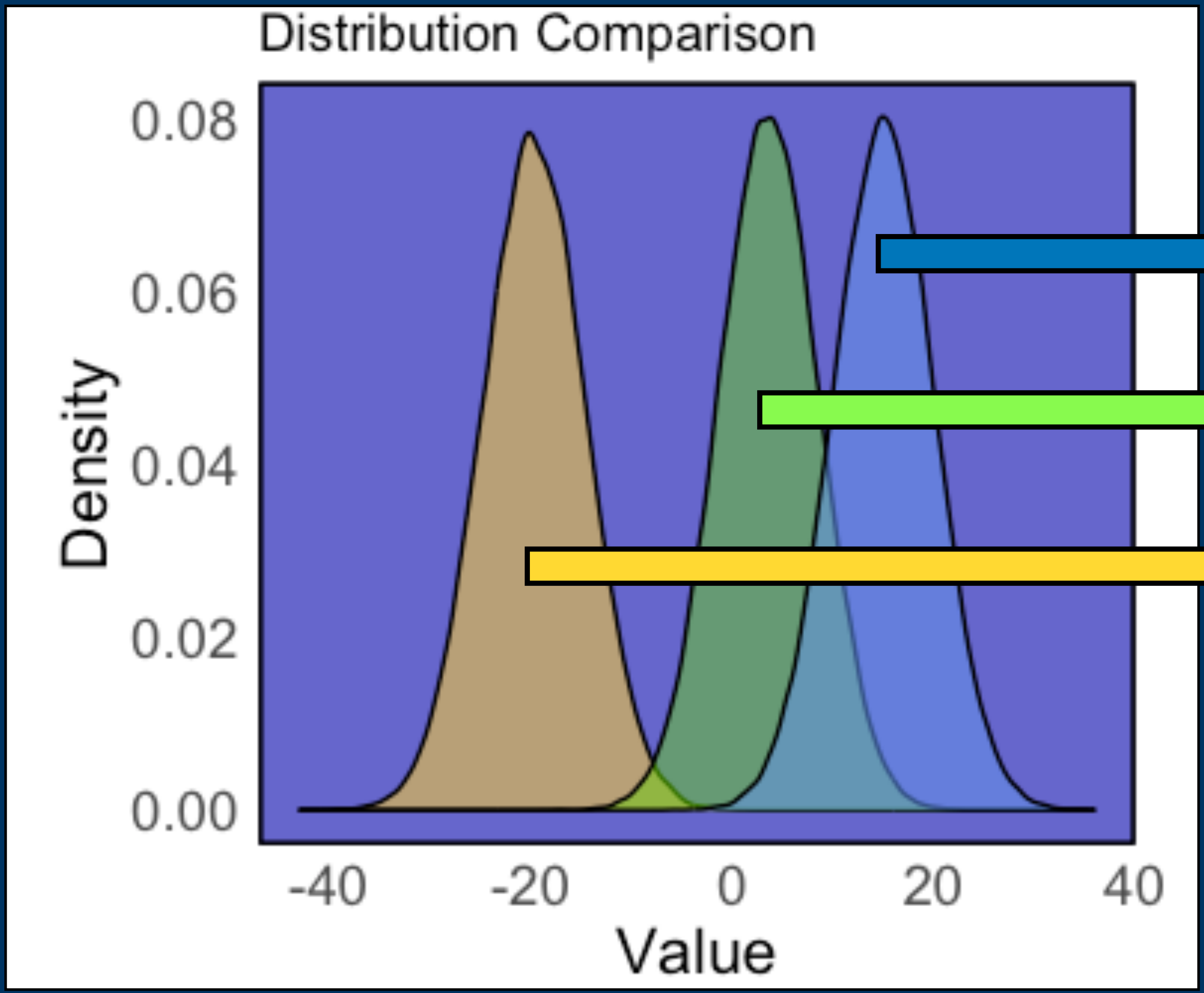
Prior N(-15, 5)

Posterior N(3.6, 4.9)

Likelihood N(-19.7, 5.1)

Case study – impact of alcohol on heart

With the bayesian approach



Prior N(-15, 5)

Posterior N(3.6, 4.9)

Likelihood N(-19.7, 5.1)

$$P(H_0 \mid x) = \frac{P(H_0) \cdot P(x \mid H_0)}{P(x)}$$

Moving forward

With the bayesian approach

In this case, we had few data (n= 21 obs). Fortunately, our 95% credible intervals include both negative and positive values, which does not allow us to conclude that drinking wine is harmful. It is important to note that this conclusion is not solely driven by theory but rather emphasizes the importance of setting priors appropriately. With more data, it becomes increasingly difficult to bias your conclusions.

Another option, which I personally prefer, is to use non-informative priors (such as uniform distributions, large standard errors, or Student's t-distributions).

Presenting a plot showing the likelihood, prior, and posterior distributions helps illustrate how the data are treated.

Moving forward

With the bayesian approach

There are plenty of distributions you can use. It will depends directly about your data.

Some ideas here:

Species cover are understood between 0 and 100%, most of the time it would be a beta distribution we need to use

Species counts refers to Poisson distribution

Multiple choices refers to multinomial (if 2, Binomial) distribution

You can have a look here: https://rdrr.io/cran/brms/man/brmsfamily.html

Thank you for your attention!

Jérémy Carlot – Postdoc MOGPA 4Oceans

14.02.24