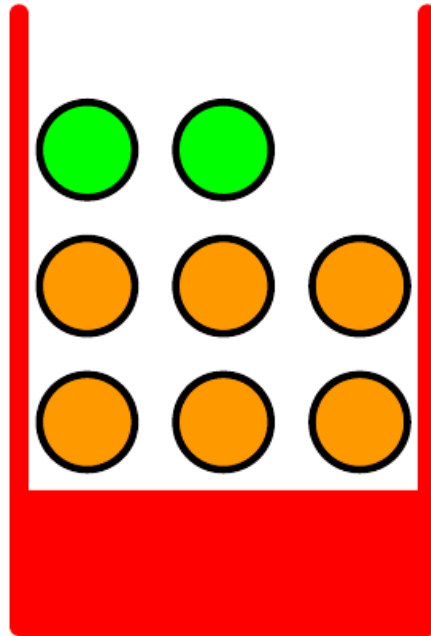
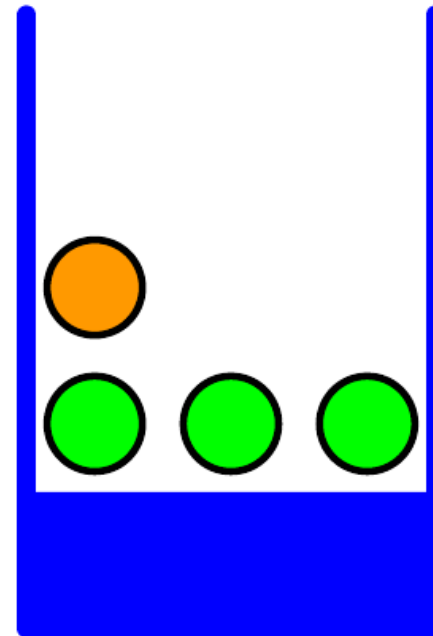


# Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box



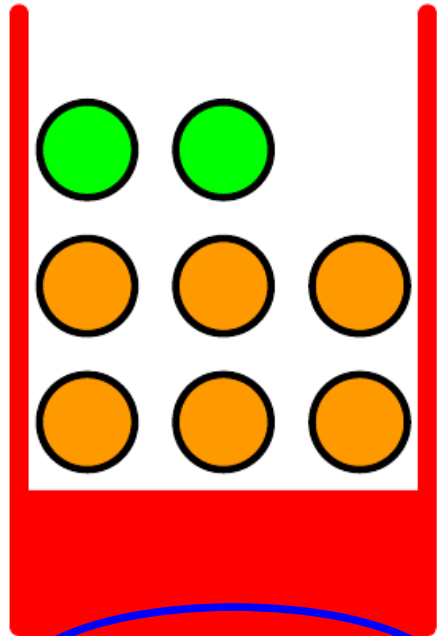
$$p(B = r) = 4/10$$



$$p(B = b) = 6/10$$

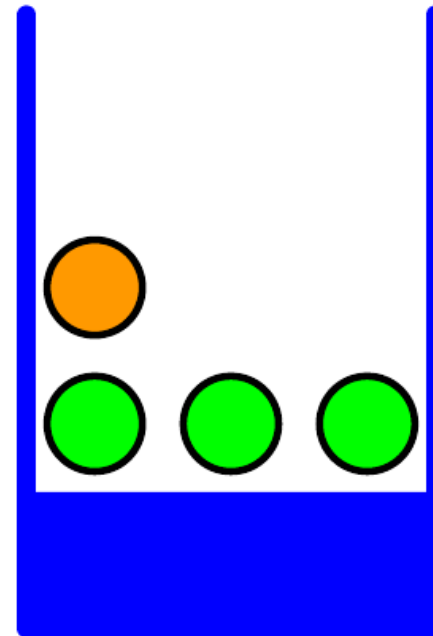
# Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box



$$p(B = r) = 4/10$$

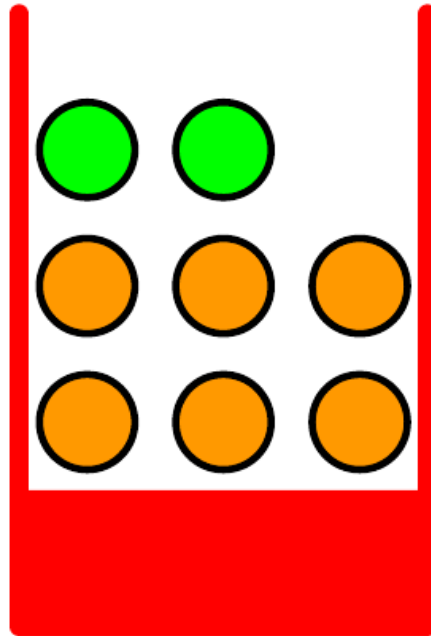
probability we  
choose red Box



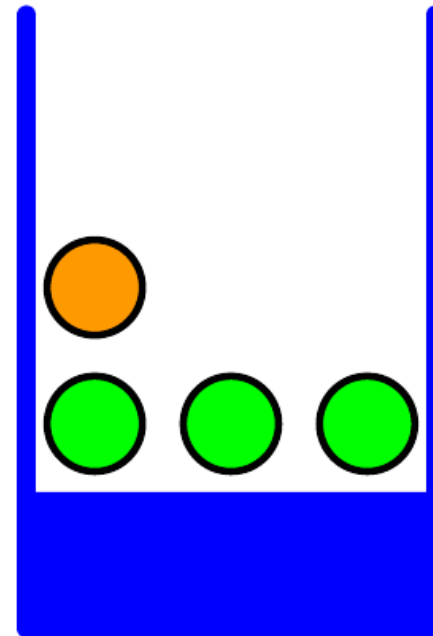
$$p(B = b) = 6/10$$

# Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box



$$p(B = r) = 4/10$$

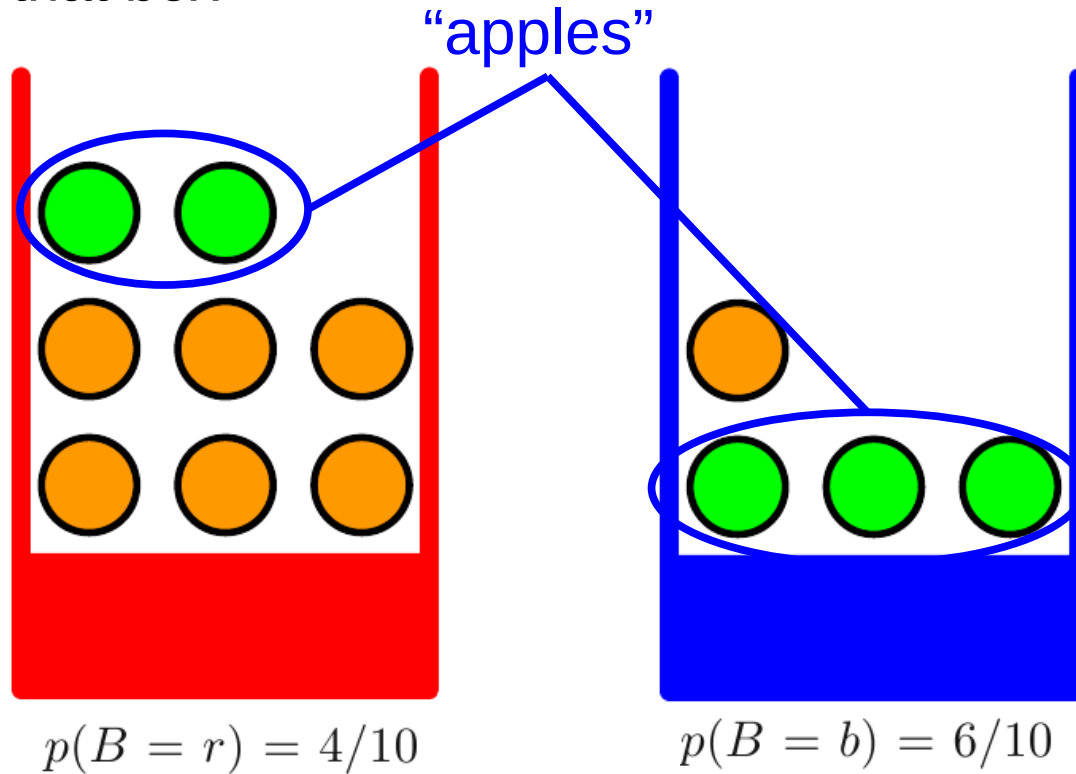


$$p(B = b) = 6/10$$

probability we  
choose blue Box

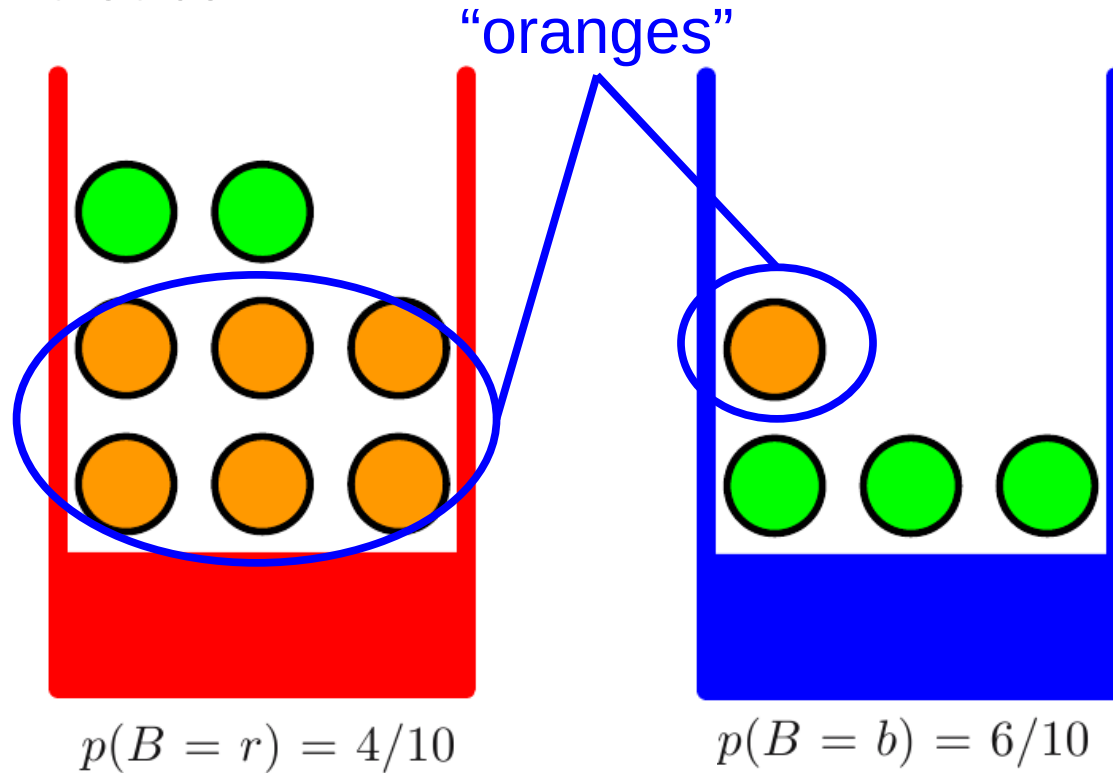
# Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box



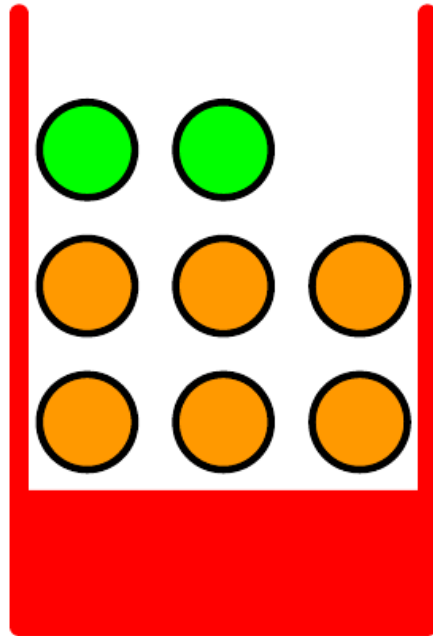
# Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box

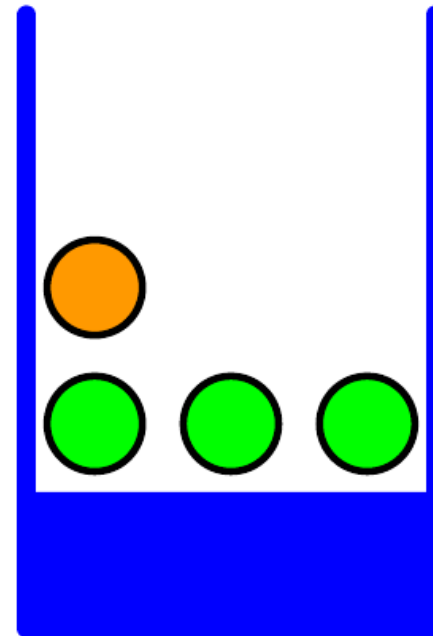


# Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box



$$p(B = r) = 4/10$$



$$p(B = b) = 6/10$$

- (1) What is the probability we get an apple?
- (2) If the fruit we end up with is an orange, what is the probability we had chosen the red box?

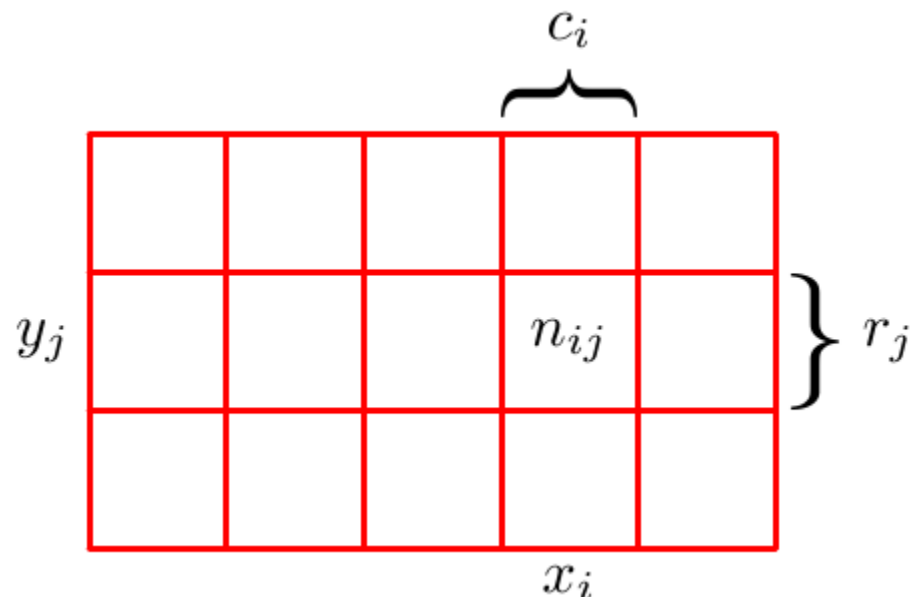
# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$



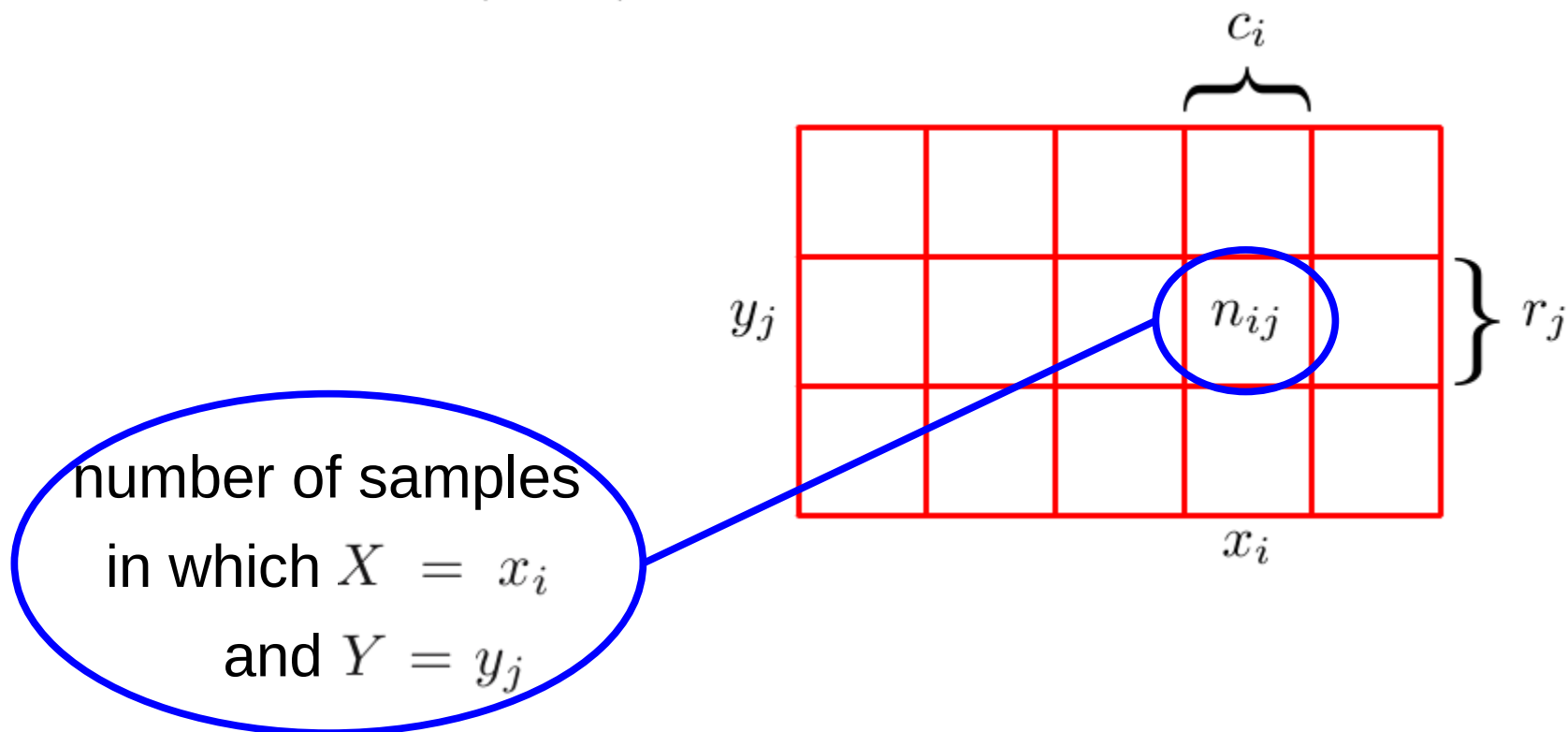
# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$





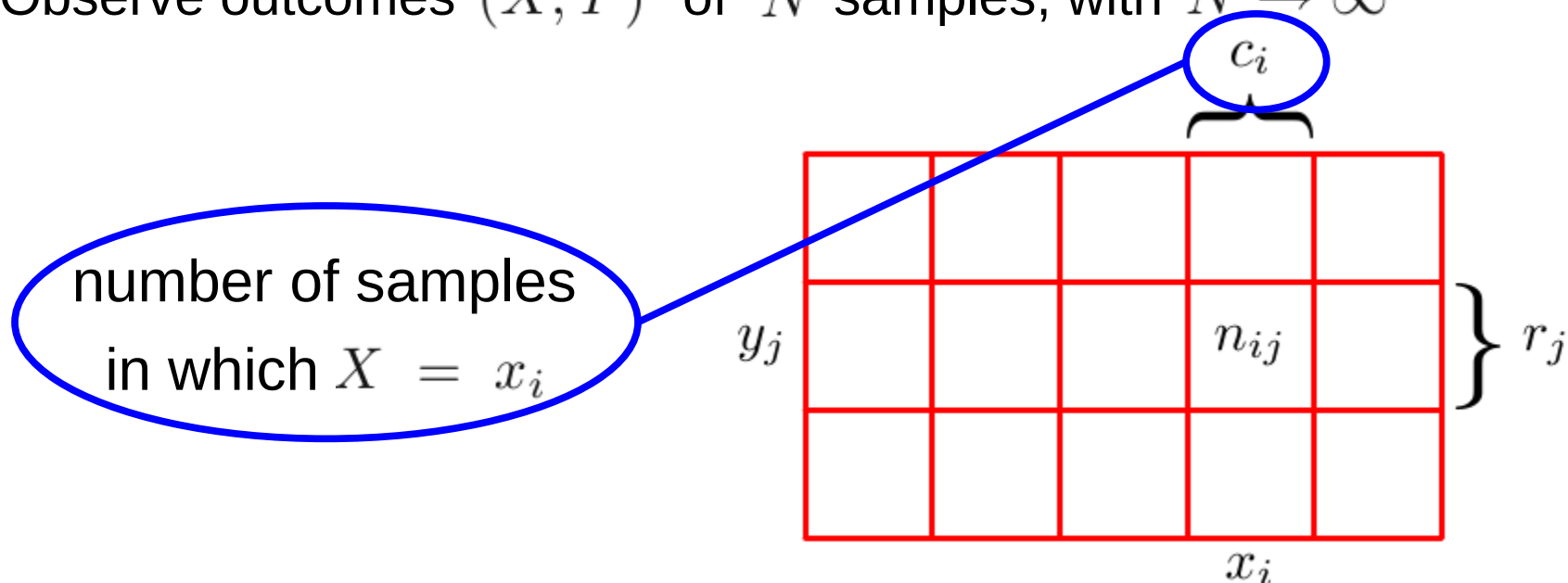
# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$



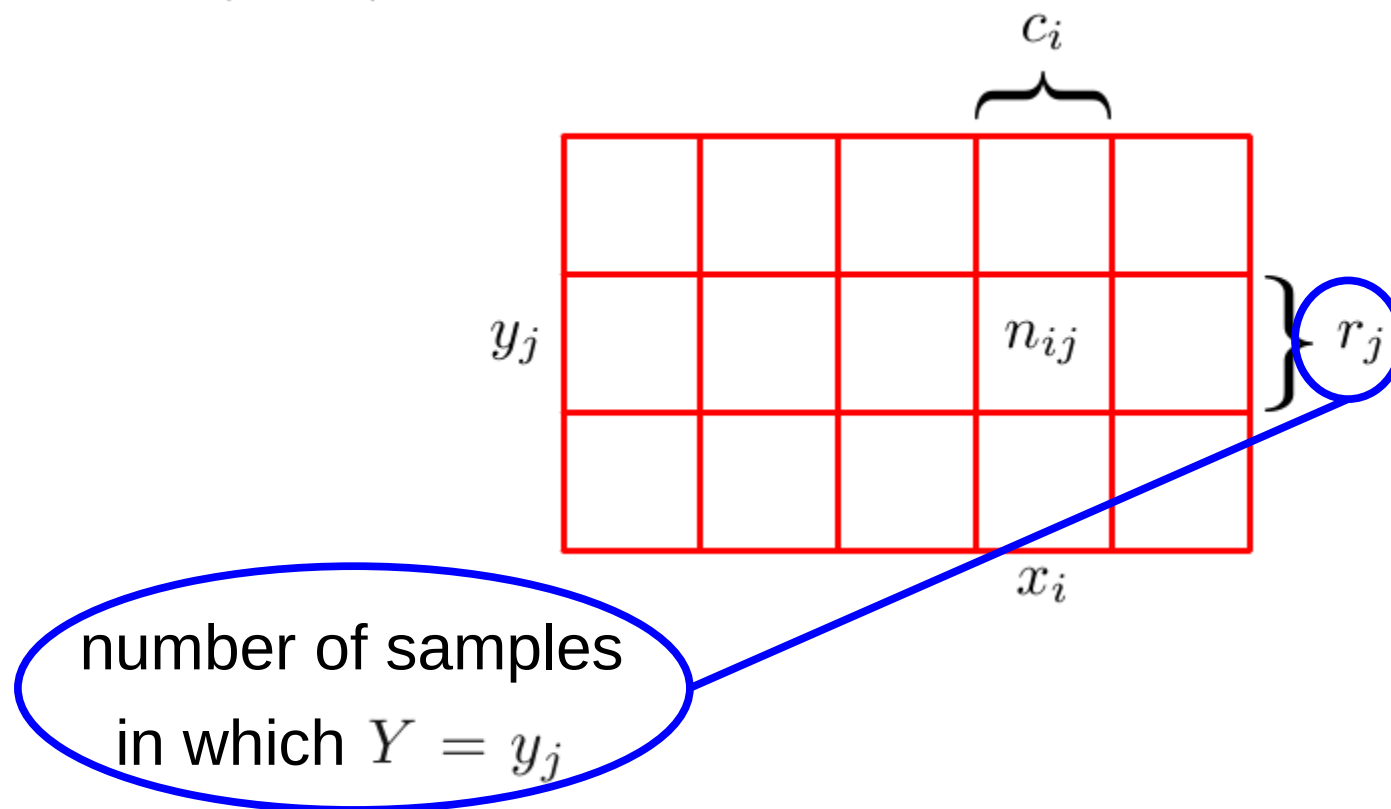
# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$



# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“joint probability”

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$

					$c_i$
$y_j$			$n_{ij}$		$r_j$
					$x_i$

# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“marginal probability”

$$p(X = x_i) = \frac{c_i}{N}$$

A 3x5 grid of cells outlined in red. A horizontal curly brace above the top row is labeled  $c_i$ . A vertical curly brace to the right of the middle row is labeled  $r_j$ . The middle row is labeled  $y_j$  on the left and  $x_i$  on the right. The cell at the intersection of the middle row and the fourth column from the left is labeled  $n_{ij}$ .

			$n_{ij}$	

# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

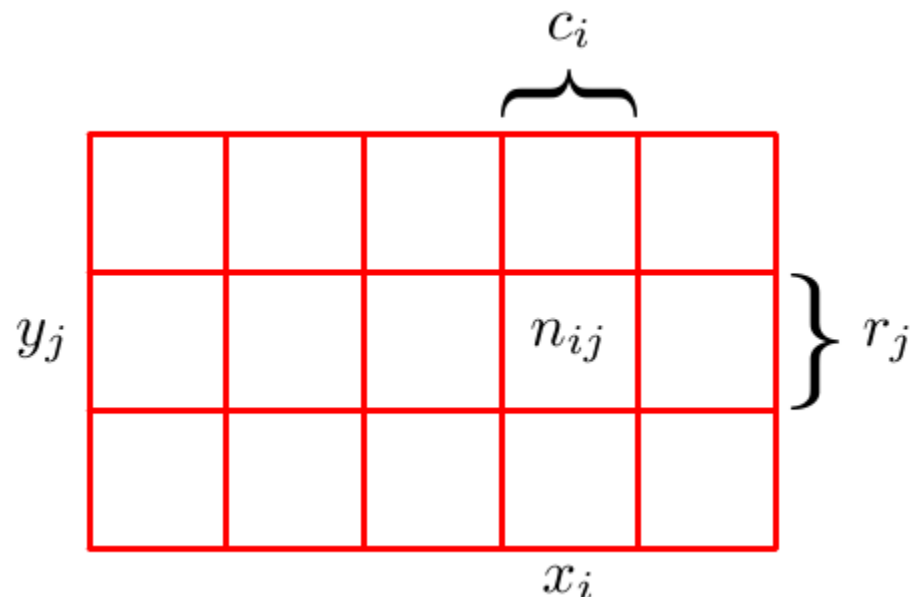
- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“marginal probability”

$$p(X = x_i) = \frac{c_i}{N}$$

$$= \frac{\sum_j n_{ij}}{N}$$

$$= \sum_{j=1}^L p(X = x_i, Y = y_j)$$



# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“marginal probability”

$$\begin{aligned} p(X = x_i) &= \frac{c_i}{N} \\ &= \frac{\sum_j n_{ij}}{N} \\ &= \sum_{j=1}^L p(X = x_i, Y = y_j) \end{aligned}$$

			$c_i$	
$y_j$			$n_{ij}$	
			$x_i$	

“sum rule”

# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“conditional probability”

$$p(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i}$$

A 3x5 grid of cells outlined in red. The top row has a brace above it labeled  $c_i$ . The middle row has a brace to its right labeled  $r_j$ . The bottom row has a label  $x_i$  centered below it. The middle row, fourth column cell contains the label  $n_{ij}$ . The label  $y_j$  is positioned to the left of the middle row.

$y_j$			$n_{ij}$	

# The two rules of probability

- Two random variables:

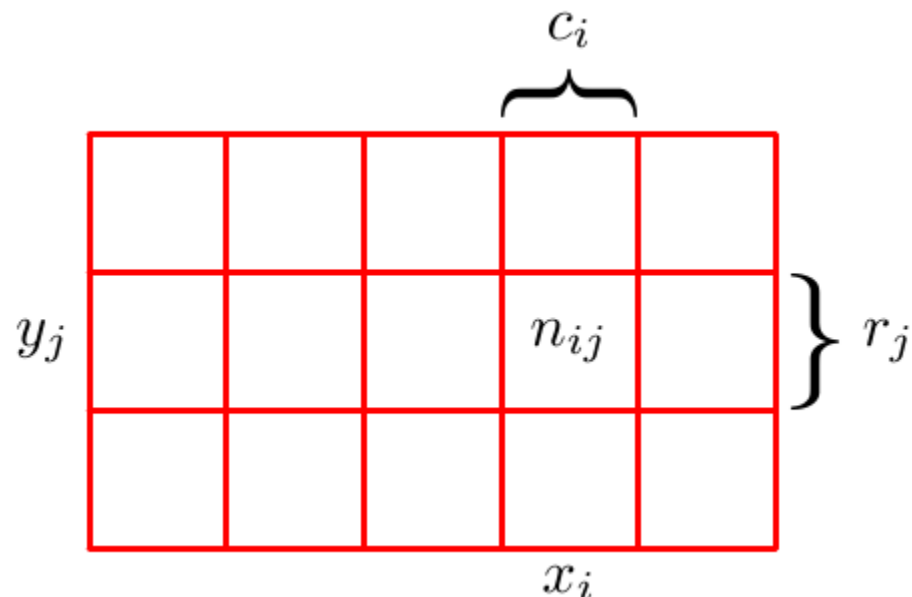
$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“joint probability”

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$





# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

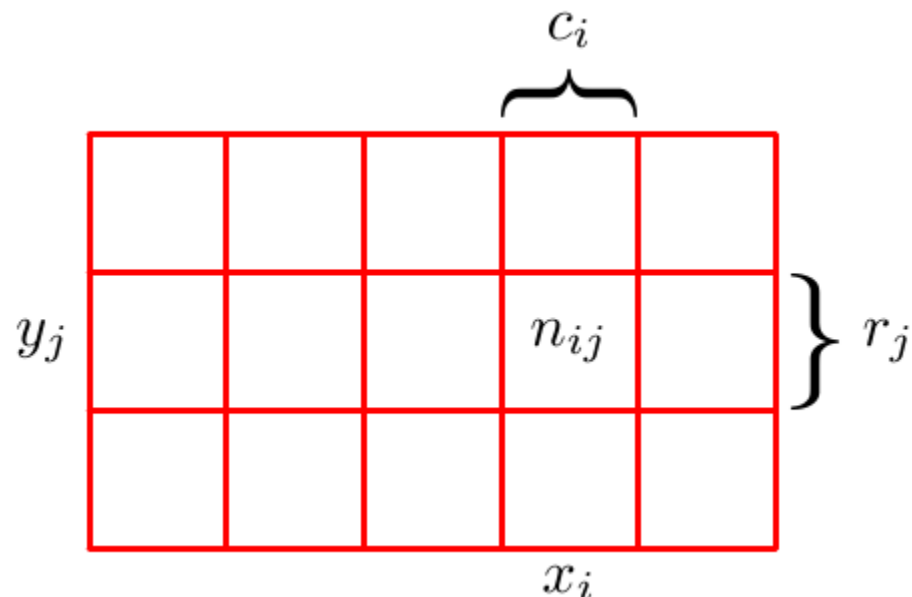
- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“joint probability”

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$

$$= \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N}$$

$$= p(Y = y_j | X = x_i) p(X = x_i)$$



# The two rules of probability

- Two random variables:

$X$  takes values  $\{x_i\}$  where  $i = 1, \dots, M$

$Y$  takes values  $\{y_j\}$  where  $j = 1, \dots, L$

- Observe outcomes  $(X, Y)$  of  $N$  samples, with  $N \rightarrow \infty$

“joint probability”

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$

$$= \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N}$$

$$= p(Y = y_j | X = x_i) p(X = x_i)$$

A 3x5 grid of cells outlined in red. The top row is labeled  $c_i$  with a brace above it. The rightmost column is labeled  $r_j$  with a brace to its right. The middle row is labeled  $y_j$  on the left. The middle column is labeled  $x_i$  below it. The cell at the intersection of the middle row and middle column contains the label  $n_{ij}$ .

					$c_i$
			$n_{ij}$		
$y_j$					$r_j$
					$x_i$

“product  
rule”

# The two rules of probability

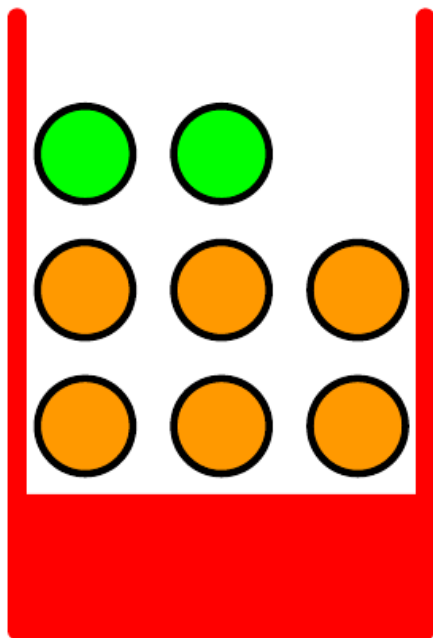
<b>sum rule</b>	$p(X) = \sum_Y p(X, Y)$
<b>product rule</b>	$p(X, Y) = p(Y X)p(X)$

- Direct result of product rule: “Bayes' rule”

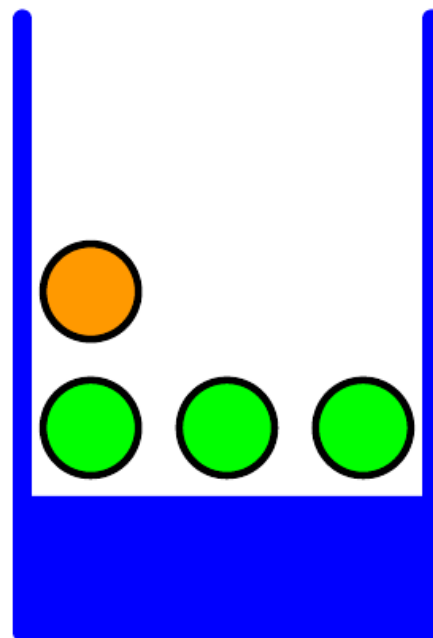
$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

# Apples and oranges

(1) What is the probability we get an apple?



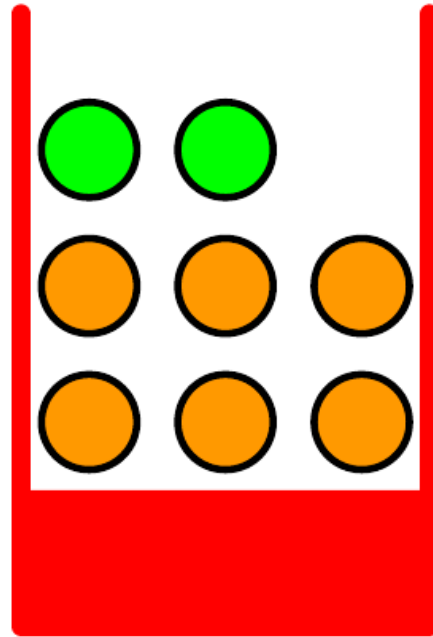
$$p(B = r) = 4/10$$



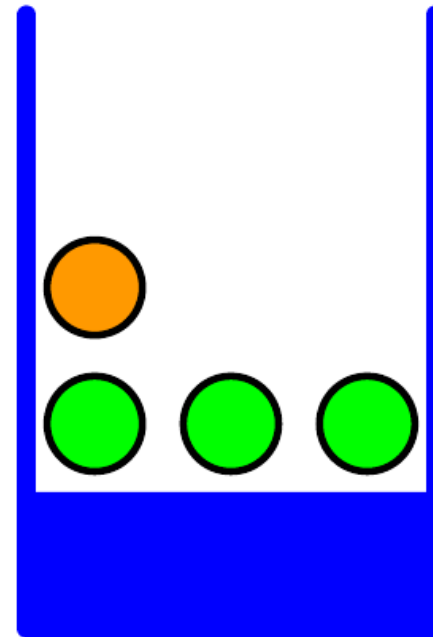
$$p(B = b) = 6/10$$

# Apples and oranges

(1) What is the probability we get an apple?



$$p(B = r) = 4/10$$

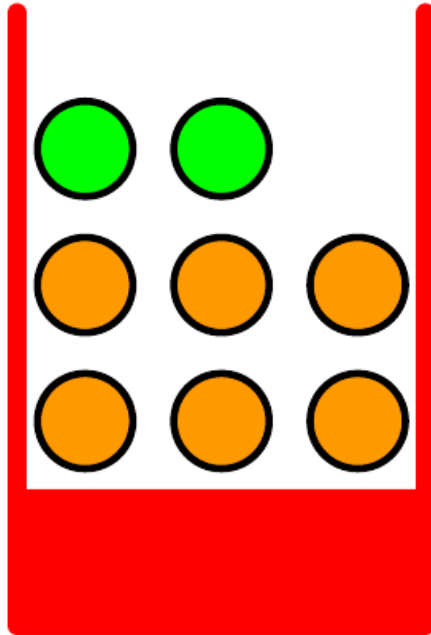


$$p(B = b) = 6/10$$

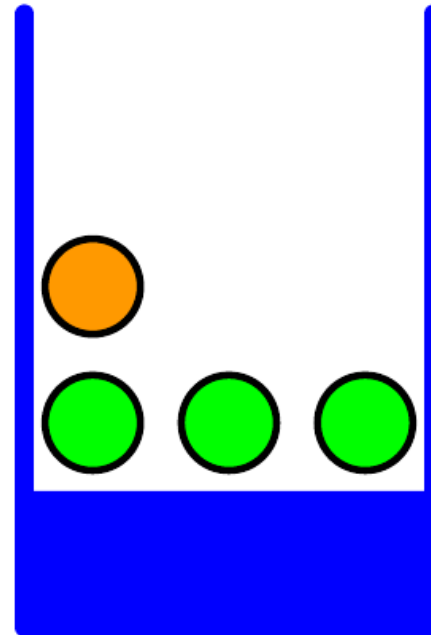
$$\begin{aligned} p(F = a) &= p(F = a|B = r)p(B = r) + p(F = a|B = b)p(B = b) \\ &= \frac{1}{4} \times \frac{4}{10} + \frac{3}{4} \times \frac{6}{10} = \frac{11}{20} \end{aligned}$$

# Apples and oranges

(2) If the fruit we end up with is an orange, what is the probability we had chosen the red box?



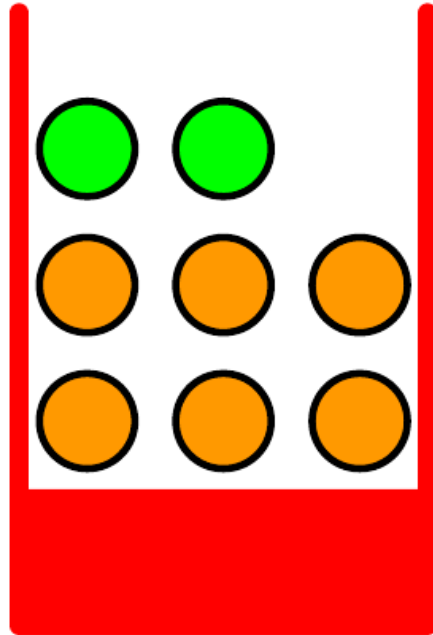
$$p(B = r) = 4/10$$



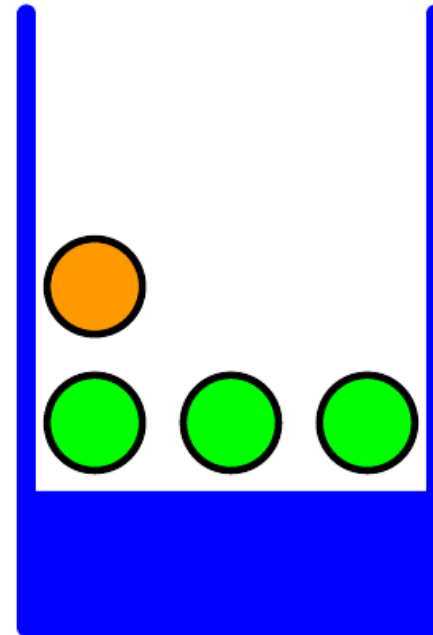
$$p(B = b) = 6/10$$

# Apples and oranges

(2) If the fruit we end up with is an orange, what is the probability we had chosen the red box?



$$p(B = r) = 4/10$$

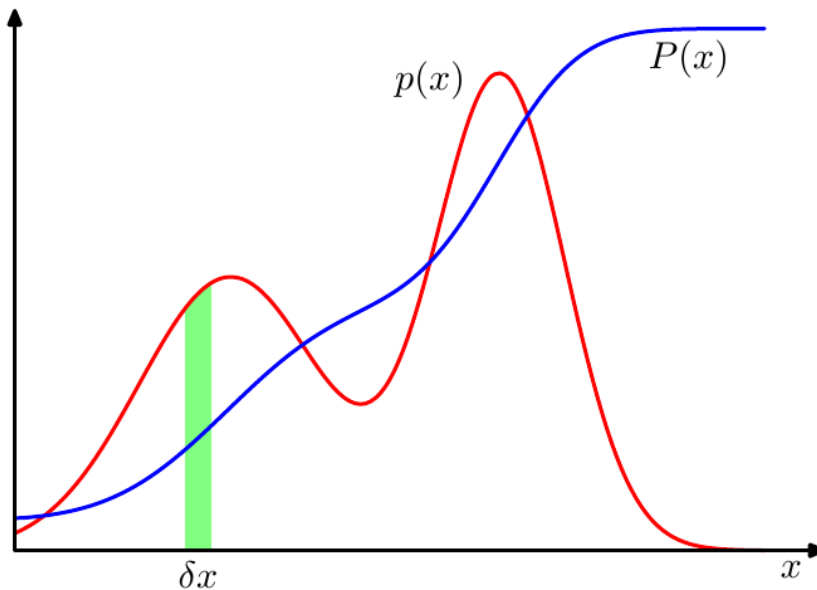


$$p(B = b) = 6/10$$

$$\begin{aligned} p(B = r | F = o) &= \frac{p(F = o | B = r)p(B = r)}{p(F = o)} \\ &= \frac{3}{4} \times \frac{4}{10} \times \frac{20}{9} = \frac{2}{3} \end{aligned}$$

# Continuous variables

- $p(x)$  is called the *probability density* over  $x$ .  
if the probability that  $x$  falls in the interval  $(x, x + \delta x)$  is given by  $p(x)\delta x$  for  $\delta x \rightarrow 0$



$$p(x) \geq 0$$
$$\int_{-\infty}^{\infty} p(x) dx = 1$$

sum rule:  $p(x) = \int p(x, y) dy$

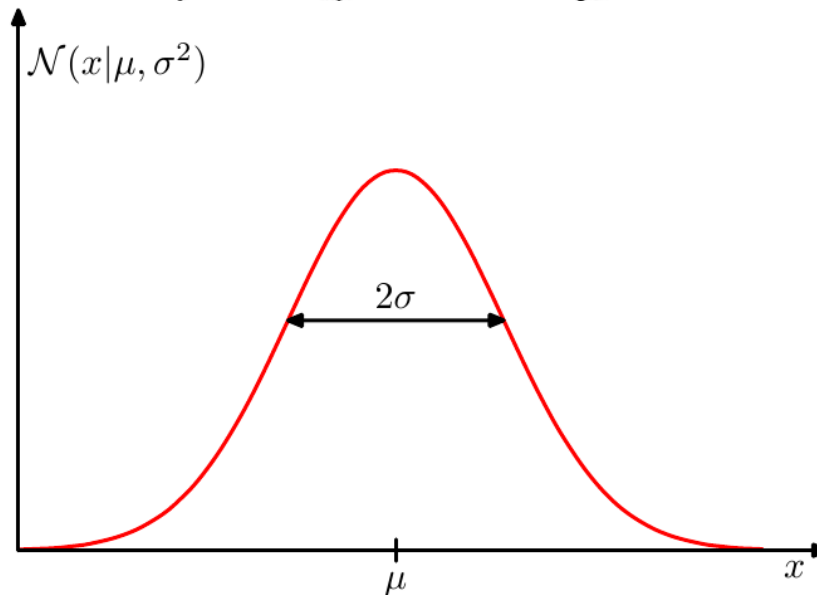
product rule:  $p(x, y) = p(y|x)p(x).$



# Continuous variables

- Example: Gaussian distribution

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\}$$



- Multivariate probability density:  $p(\mathbf{x}) = p(x_1, \dots, x_D)$

$$p(\mathbf{x}) \geq 0$$

$$\int p(\mathbf{x}) d\mathbf{x} = 1$$