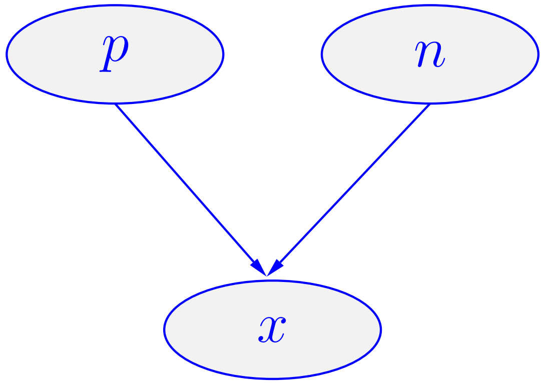
n independent Bernoulli processes

General case

n independent Bernoulli processes

General case

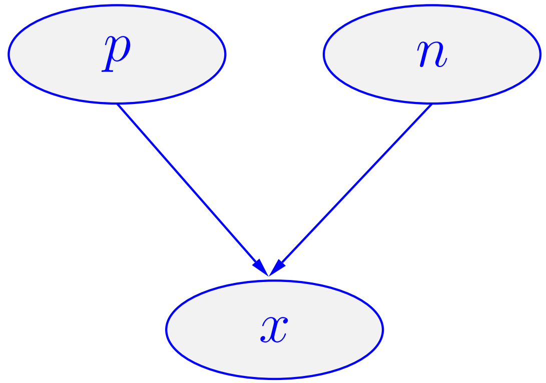
Model



n independent Bernoulli processes

General case

Model



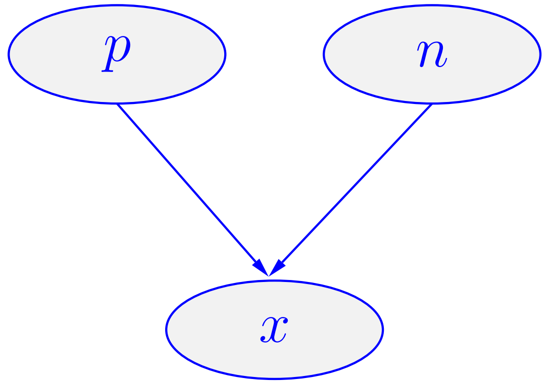
Joint pdf (omitting background condition I):

$$f(x, p, n) \;=\; f(x \,|\, p, n)$$

n independent Bernoulli processes

General case

Model



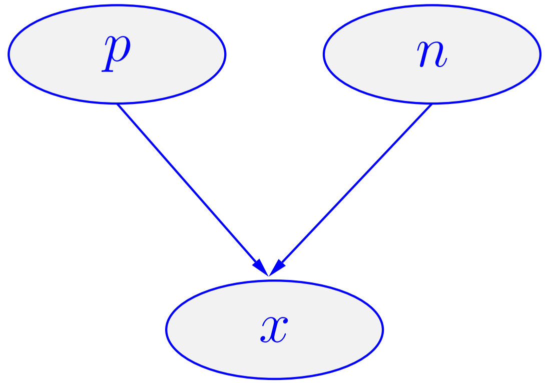
Joint pdf (omitting background condition I):

$$f(x, p, n) \;=\; f(x \,|\, p, n) \cdot f(p, n)$$

n independent Bernoulli processes

General case

Model



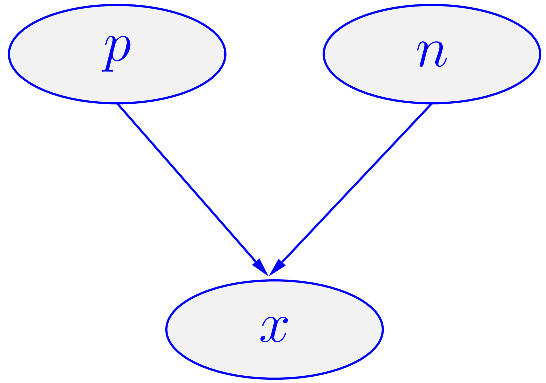
Joint pdf (omitting background condition I):

$$
\begin{aligned}
f(x, p, n) &= f(x \mid p, n) \cdot f(p, n) \\
&= f(x \mid p, n) \cdot f(p \mid n) \cdot f(n)
\end{aligned}
$$

n independent Bernoulli processes

General case

Model



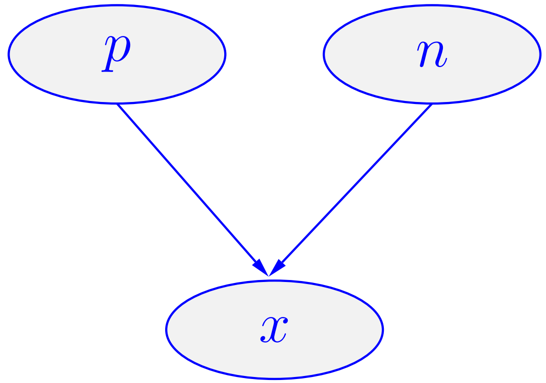
Joint pdf (omitting background condition I):

$$
\begin{aligned}
f(x, p, n) &= f(x \mid p, n) \cdot f(p, n) \\
&= f(x \mid p, n) \cdot f(p \mid n) \cdot f(n) \\
&= f(x \mid p, n) \cdot f(n \mid p) \cdot f(p)
\end{aligned}
$$

n independent Bernoulli processes

General case

Model



Joint pdf (omitting background condition I):

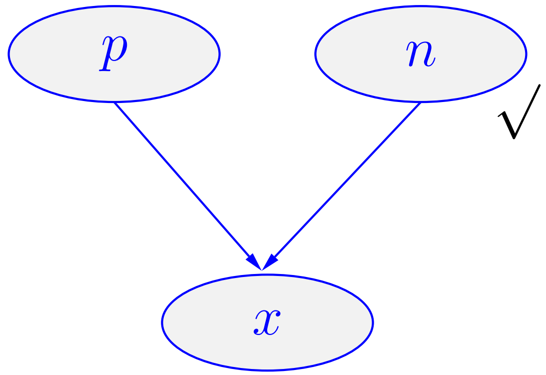
$$
\begin{aligned}
f(x, p, n) &= f(x \,|\, p, n) \cdot f(p, n) \\
&= f(x \,|\, p, n) \cdot f(p \,|\, n) \cdot f(n) \\
&= f(x \,|\, p, n) \cdot f(n \,|\, p) \cdot f(p) \\
&= f(x \,|\, p, n) \cdot f(p) \cdot f(n)
\end{aligned}
$$

(n and p are independent)

n independent Bernoulli processes

Usual case → n fixed (for the moment)

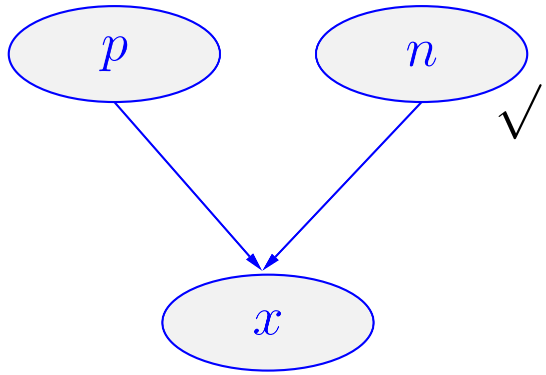
Model



n independent Bernoulli processes

Usual case → n fixed (for the moment)

Model



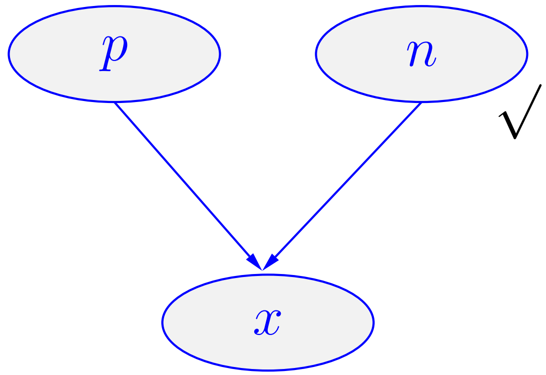
Joint pdf

$$f(x, p \,|\, n) \;=\; f(x \,|\, p, n) \cdot f(p)$$

n independent Bernoulli processes

Usual case → n fixed (for the moment)

Model



Joint pdf

$$f(x, p \,|\, n) \;=\; f(x \,|\, p, n) \cdot f(p)$$

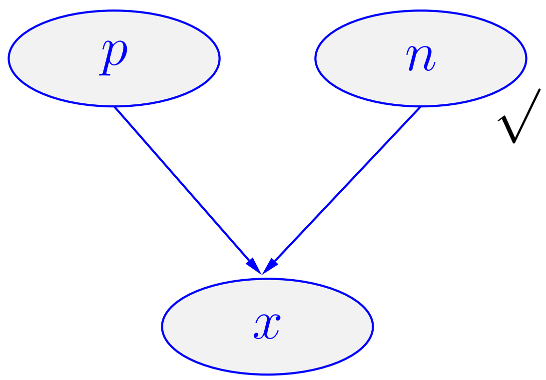
Typical problems

- p is assumed → interested in $f(x \,|\, n, p)$
  → well known binomial;

n independent Bernoulli processes

Usual case → n fixed (for the moment)

Model



Joint pdf

$$f(x, p \mid n) \;=\; f(x \mid p, n) \cdot f(p)$$

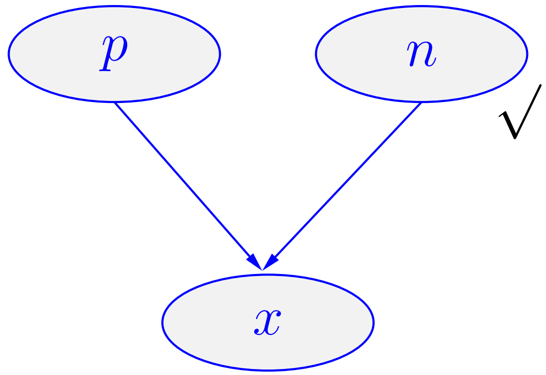
Typical problems

- p is assumed → interested in $f(x \mid n, p)$
  → well known binomial;
- x is assumed (‘observed’)

n independent Bernoulli processes

Usual case → n fixed (for the moment)

Model



Joint pdf

$$f(x, p \mid n) \;=\; f(x \mid p, n) \cdot f(p)$$

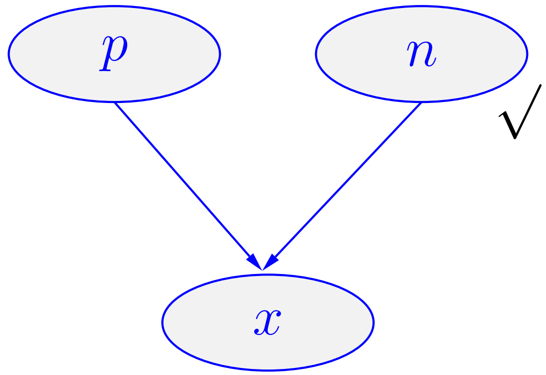
Typical problems

- p is assumed → interested in $f(x \mid n, p)$
  → well known binomial;
- x is assumed ('observed') → $f(p \mid n, x)$

n independent Bernoulli processes

Usual case $\rightarrow$ n fixed (for the moment)

Model



Joint pdf

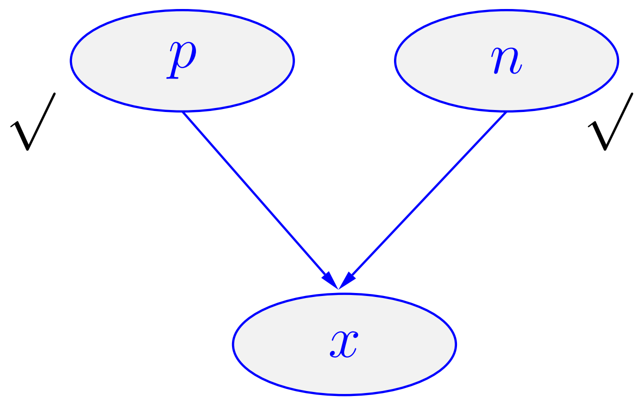
$$f(x, p \,|\, n) \;=\; f(x \,|\, p, n) \cdot f(p)$$

Typical problems

- p is assumed $\rightarrow$ interested in $f(x \,|\, n, p)$
  $\rightarrow$ well known binomial;
- x is assumed ('observed') $\rightarrow$ $f(p \,|\, n, x)$:
  $\rightarrow$?

n independent Bernoulli processes

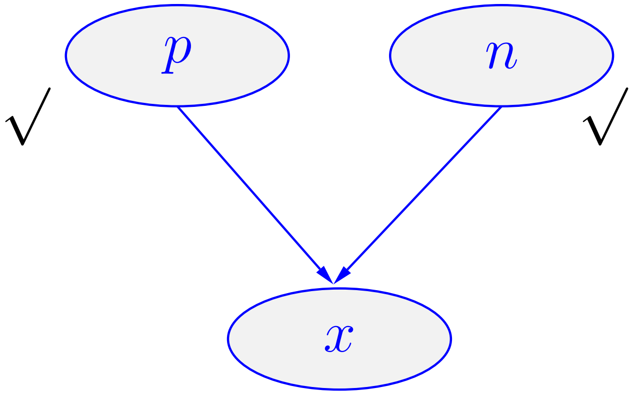
Graphical models of the typical problems



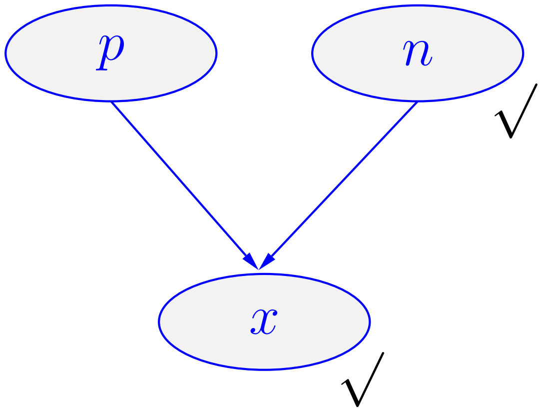
$$\rightarrow f(x \mid n, p)$$

n independent Bernoulli processes

Graphical models of the typical problems



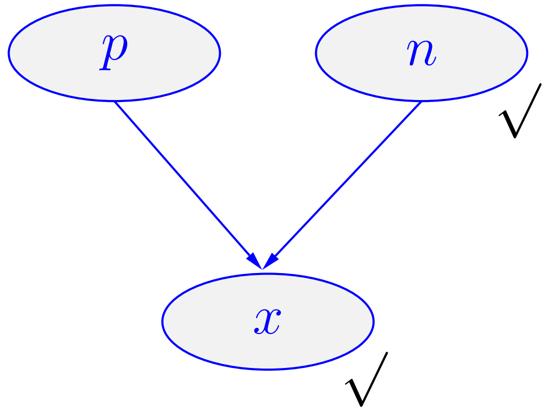
$\rightarrow\ f(x\,|\,n,p)$



$\rightarrow\ f(p\,|\,n,x)$

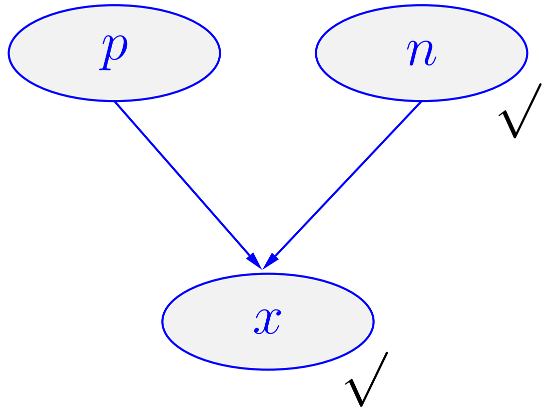
n independent Bernoulli processes

Inferring p



n independent Bernoulli processes

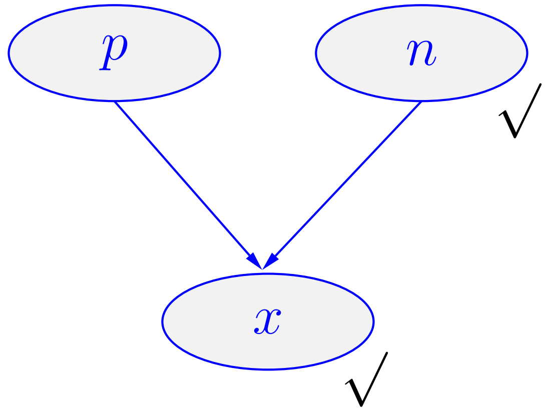
Inferring p



$$f(p\,|\,x,n) \;=\; \frac{f(p,x\,|\,n)}{f(x\,|\,n)}$$

n independent Bernoulli processes

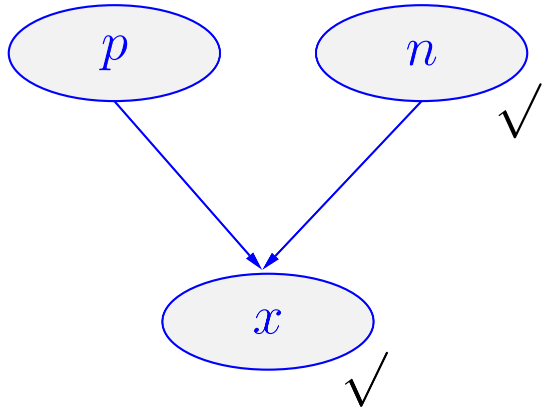
Inferring p



$$
\begin{aligned}
f(p \mid x, n) &= \frac{f(p, x \mid n)}{f(x \mid n)} \\
&= \frac{f(x \mid n, p) \cdot f_0(p)}{f(x \mid n)}
\end{aligned}
$$

n independent Bernoulli processes

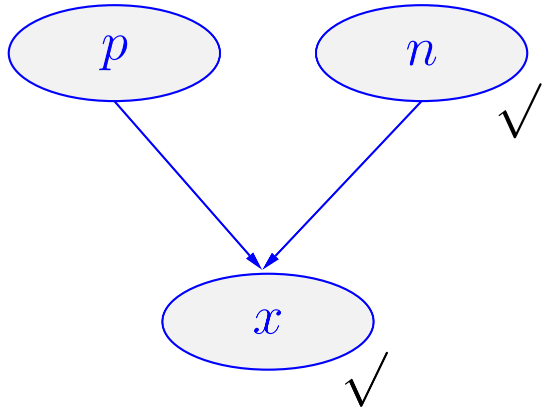
Inferring p



$$
\begin{aligned}
f(p \mid x, n) &= \frac{f(p, x \mid n)}{f(x \mid n)} \\
&= \frac{f(x \mid n, p) \cdot f_0(p)}{f(x \mid n)} \\
&= \frac{f(x \mid n, p) \cdot f_0(p)}{\int_0^1 f(x \mid n, p) \cdot f_0(p)\, dp}
\end{aligned}
$$

n independent Bernoulli processes

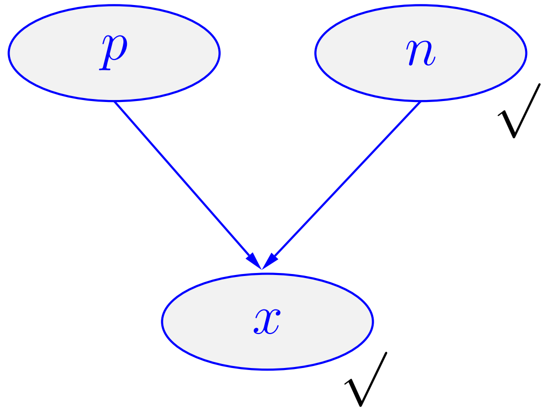
Inferring p



$$
\begin{aligned}
f(p\,|\,x,n) &= \frac{f(p,x\,|\,n)}{f(x\,|\,n)} \\
&= \frac{f(x\,|\,n,p)\cdot f_0(p)}{f(x\,|\,n)} \\
&= \frac{f(x\,|\,n,p)\cdot f_0(p)}{\int_0^1 f(x\,|\,n,p)\cdot f_0(p)\,dp} \\
&\propto f(x\,|\,n,p)\cdot f_0(p)
\end{aligned}
$$

n independent Bernoulli processes

Inferring p



$$
\begin{aligned}
f(p \mid x, n) &= \frac{f(p, x \mid n)}{f(x \mid n)} \\
&= \frac{f(x \mid n, p) \cdot f_0(p)}{f(x \mid n)} \\
&= \frac{f(x \mid n, p) \cdot f_0(p)}{\int_0^1 f(x \mid n, p) \cdot f_0(p)\, dp} \\
&\propto f(x \mid n, p) \cdot f_0(p)
\end{aligned}
$$

(denominator just normalization!)

Inferring "Bernoulli's p"

We just need to make explicit $f(x \,|\, n, p)$:

$$f(x \,|\, n, p) \;=\; \binom{n}{x}\, p^x\, (1-p)^{n-x}$$

Inferring "Bernoulli's p"

We just need to make explicit $f(x \mid n, p)$:

$$f(x \mid n, p) \;=\; \binom{n}{x} p^x (1-p)^{n-x} \;=\; \frac{n!}{(n-x)!\,x!}\, p^x (1-p)^{n-x}$$

Inferring "Bernoulli's p"

We just need to make explicit $f(x\,|\,n,p)$:

$$f(x\,|\,n,p) \;=\; \binom{n}{x}\,p^x\,(1-p)^{n-x} \;=\; \frac{n!}{(n-x)!\,x!}\,p^x\,(1-p)^{n-x}$$

We get then, including normalization:

$$f(p\,|\,x,n) \;=\; \frac{\frac{n!}{(n-x)!\,x!}\,p^x\,(1-p)^{n-x}\,f_\circ(p)}{\int_0^1 \frac{n!}{(n-x)!\,x!}\,p^x\,(1-p)^{n-x}\,f_\circ(p)\,\mathrm{d}p}$$

Inferring “Bernoulli's p”

We just need to make explicit $f(x\,|\,n,p)$:

$$f(x\,|\,n,p) \;=\; \binom{n}{x}\, p^x\,(1-p)^{n-x} \;=\; \frac{n!}{(n-x)!\,x!}\, p^x\,(1-p)^{n-x}$$

We get then, including normalization:

$$\begin{aligned} f(p\,|\,x,n) &= \frac{\frac{n!}{(n-x)!\,x!}\, p^x\,(1-p)^{n-x}\, f_\circ(p)}{\int_0^1 \frac{n!}{(n-x)!\,x!}\, p^x\,(1-p)^{n-x}\, f_\circ(p)\,\mathrm{d}p} \\ &= \frac{p^x\,(1-p)^{n-x}\, f_\circ(p)}{\int_0^1 p^x\,(1-p)^{n-x}\, f_\circ(p)\,\mathrm{d}p} \end{aligned}$$

Inferring "Bernoulli's p"

We just need to make explicit $f(x \,|\, n, p)$:

$$f(x \,|\, n, p) \;=\; \binom{n}{x} p^x (1-p)^{n-x} \;=\; \frac{n!}{(n-x)!\,x!}\, p^x (1-p)^{n-x}$$

We get then, including normalization:

$$f(p \,|\, x, n) \;=\; \frac{\frac{n!}{(n-x)!\,x!}\, p^x (1-p)^{n-x} f_\circ(p)}{\int_0^1 \frac{n!}{(n-x)!\,x!}\, p^x (1-p)^{n-x} f_\circ(p)\,\mathrm{d}p}$$

$$=\; \frac{p^x (1-p)^{n-x} f_\circ(p)}{\int_0^1 p^x (1-p)^{n-x} f_\circ(p)\,\mathrm{d}p}$$

(The binomial coefficient is irrelevant, not depending on p)

Inferring "Bernoulli's p"

$$f(p \,|\, x, n) \;=\; \frac{p^x\,(1-p)^{n-x}\,f_\circ(p)}{\int_0^1 p^x\,(1-p)^{n-x}\,f_\circ(p)\,\mathrm{d}p}$$

Inferring "Bernoulli's p"

$$f(p\,|\,x,n) \;=\; \frac{p^x\,(1-p)^{n-x}\,f_\circ(p)}{\int_0^1 p^x\,(1-p)^{n-x}\,f_\circ(p)\,\mathrm{d}p}$$

For teaching purposes we start from a **uniform prior**, i.e. $\boldsymbol{f_\circ(p)} = 1$:

$$f(p\,|\,x,n) \;=\; \frac{p^x\,(1-p)^{n-x}}{\int_0^1 p^x\,(1-p)^{n-x}\,\mathrm{d}p}$$

Inferring "Bernoulli's p"

$$f(p\,|\,x,n) \;=\; \frac{p^x\,(1-p)^{n-x}\,f_\circ(p)}{\int_0^1 p^x\,(1-p)^{n-x}\,f_\circ(p)\,\mathrm{d}p}$$

For teaching purposes we start from a **uniform prior**, i.e. $\boldsymbol{f_\circ(p) = 1}$:

$$f(p\,|\,x,n) \;=\; \frac{p^x\,(1-p)^{n-x}}{\int_0^1 p^x\,(1-p)^{n-x}\,\mathrm{d}p}$$

- The integral at the denominator is the special function "β" (also defined for real values of x and n).

Inferring "Bernoulli's p"

$$f(p\,|\,x,n) \;=\; \frac{p^x\,(1-p)^{n-x}\,f_\circ(p)}{\int_0^1 p^x\,(1-p)^{n-x}\,f_\circ(p)\,\mathrm{d}p}$$

For teaching purposes we start from a **uniform prior**, i.e. $\boldsymbol{f_\circ(p)} = 1$:

$$f(p\,|\,x,n) \;=\; \frac{p^x\,(1-p)^{n-x}}{\int_0^1 p^x\,(1-p)^{n-x}\,\mathrm{d}p}$$

- The integral at the denominator is the special function "β" (also defined for real values of x and n).
- In our case these two numbers are integer and the integral becomes equal to

$$\frac{x!\,(n-x)!}{(n+1)!}$$

Inferring "Bernoulli's p"

Solution for uniform prior (think to **Bayes' billard**)

$$f(p\,|\,x,n) \quad = \quad \frac{(n+1)!}{x!\,(n-x)!}\,p^x\,(1-p)^{n-x}$$

Inferring "Bernoulli's p"

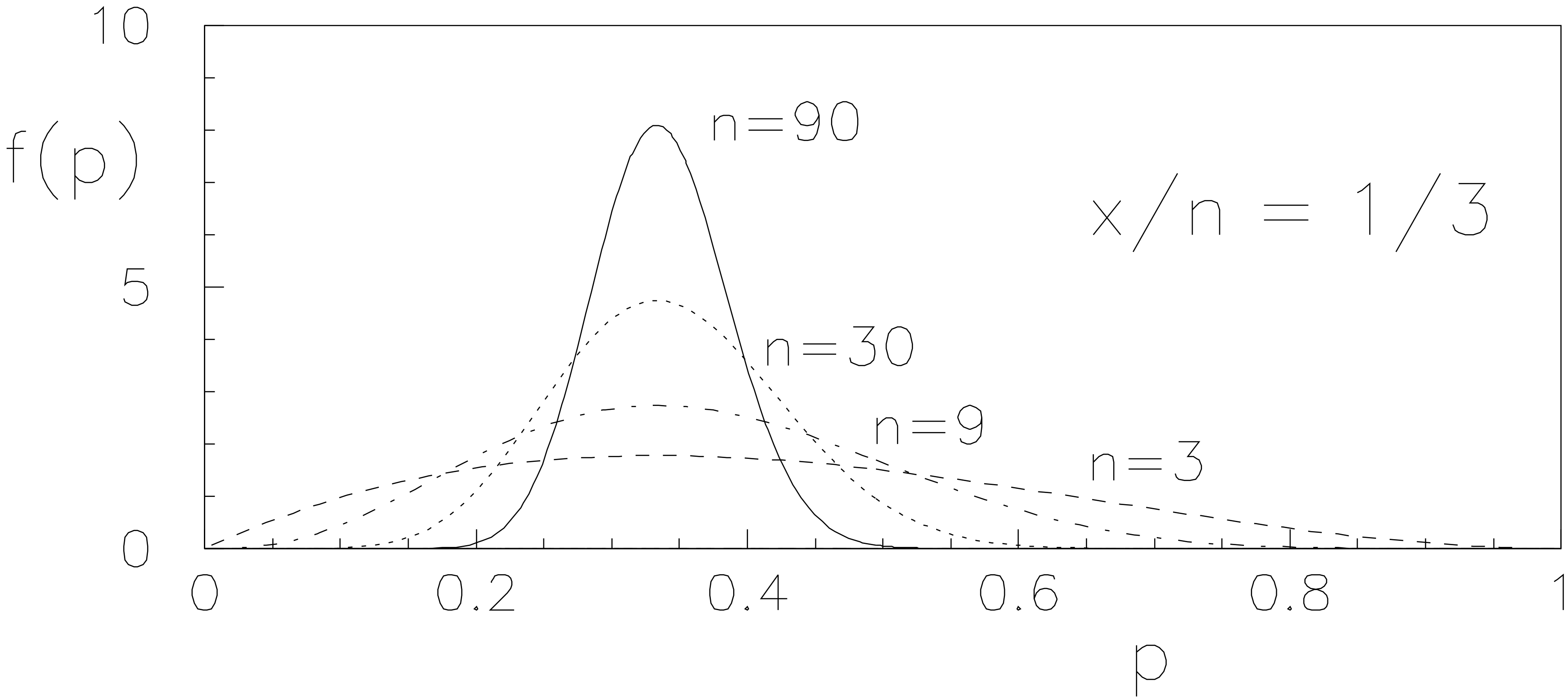
Solution for uniform prior (think to **Bayes' billard**)

$$
\begin{aligned}
f(p \,|\, x, n) &= \frac{(n+1)!}{x!\,(n-x)!}\, p^x\,(1-p)^{n-x} \\
&= (n+1)\cdot\frac{n!}{x!\,(n-x)!}\, p^x\,(1-p)^{n-x}
\end{aligned}
$$

Inferring "Bernoulli's p"

Solution for uniform prior (think to **Bayes' billard**)

$$
\begin{aligned}
f(p\,|\,x,n) &= \frac{(n+1)!}{x!\,(n-x)!}\,p^x\,(1-p)^{n-x} \\
&= (n+1)\cdot\frac{n!}{x!\,(n-x)!}\,p^x\,(1-p)^{n-x}
\end{aligned}
$$



Inferring "Bernoulli's p"

Summaries of the **posterior distribution**

$$p_m = \text{mode}(p) \quad = \quad \frac{x}{n}$$

Inferring "Bernoulli's p"

Summaries of the **posterior distribution**

$$p_m = \text{mode}(p) = \frac{x}{n}$$

$$E(p) = \frac{x+1}{n+2}$$

Inferring "Bernoulli's p"

Summaries of the **posterior distribution**

$$p_m = \text{mode}(p) = \frac{x}{n}$$

$$E(p) = \frac{x+1}{n+2}$$

"recursive Laplace formula"

("Laplace's rule of succession")

Inferring "Bernoulli's p"

Summaries of the **posterior distribution**

$$p_m = \text{mode}(p) = \frac{x}{n}$$

$$E(p) = \frac{x+1}{n+2}$$

"recursive Laplace formula"

("Laplace's rule of succession")

$$\begin{aligned}\text{Var}(p) &= \frac{(x+1)(n-x+1)}{(n+3)(n+2)^2} \\ &= \frac{x+1}{n+2}\left(\frac{n+2}{n+2} - \frac{x+1}{n+2}\right)\frac{1}{n+3} \\ &= E(p)\,(1-E(p))\,\frac{1}{n+3}\end{aligned}$$

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".

  [$^{(*)}$ "First" does not imply time order, but just order in usage.]

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p,

Inferring the “Bernoulli's p”

About the meaning of E(p)

- We have used the “first”$^{(*)}$ n trials to learn about “p”.
  [$^{(*)}$ “First” does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i \,|\, p) = p$$

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i \,|\, p) = p$$

- But since we are uncertain about it, we have to take into account all possible values, weighing them with
  our degree of belief.

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i \,|\, p) = p$$

- But since we are uncertain about it, we have to take into account all possible values, weighing them with our degree of belief.

$$P(E_{i>n} \,|\, x, n) \;=\; \int_0^1 P(E_i \,|\, p) \cdot f(p \,|\, x, n)\, \mathrm{d}p$$

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i \,|\, p) = p$$

- But since we are uncertain about it, we have to take into account all possible values, weighing them with our degree of belief.

$$\begin{aligned} P(E_{i>n} \,|\, x, n) &= \int_0^1 P(E_i \,|\, p) \cdot f(p \,|\, x, n)\, \mathrm{d}p \\ &= \int_0^1 p \cdot f(p \,|\, x, n)\, \mathrm{d}p \end{aligned}$$

Inferring the "Bernoulli's p"

About the meaning of E(p)

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i \,|\, p) = p$$

- But since we are uncertain about it, we have to take into account all possible values, weighing them with our degree of belief.

$$\begin{aligned} P(E_{i>n} \,|\, x, n) &= \int_0^1 P(E_i \,|\, p) \cdot f(p \,|\, x, n)\, \mathrm{d}p \\ &= \int_0^1 p \cdot f(p \,|\, x, n)\, \mathrm{d}p = \mathrm{E}(p) \quad (!!) \end{aligned}$$

Inferring the "Bernoulli's p"

About the meaning of $\mathrm{E}(p)$

- We have used the "first"$^{(*)}$ n trials to learn about "p".
  [$^{(*)}$ "First" does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i\,|\,p) = p$$

- But since we are uncertain about it, we have to take into account all possible values, weighing them with our degree of belief.

$$\begin{aligned} P(E_{i>n}\,|\,x,n) &= \int_0^1 P(E_i\,|\,p)\cdot f(p\,|\,x,n)\,\mathrm{d}p \\ &= \int_0^1 p\cdot f(p\,|\,x,n)\,\mathrm{d}p = \mathrm{E}(p) \quad (!!) \end{aligned}$$

$\mathrm{E}(p)$ (and not the mode!) is the probability of every 'future' event which is believed to have the **same p** of the 'previous' ones.

Inferring the “Bernoulli’s p”

About the meaning of $\mathrm{E}(p)$

- We have used the “first”$^{(*)}$ n trials to learn about “p”.
  [$^{(*)}$ “First” does not imply time order, but just order in usage.]
- What will be the probability of other trials?

$$P(E_{i>n}) = ??$$

- If we were sure about p, then p would be our probability:

$$P(E_i \,|\, p) = p$$

- But since we are uncertain about it, we have to take into account all possible values, weighing them with our degree of belief.

$$\begin{aligned} P(E_{i>n} \,|\, x, n) &= \int_0^1 P(E_i \,|\, p) \cdot f(p \,|\, x, n) \, \mathrm{d}p \\ &= \int_0^1 p \cdot f(p \,|\, x, n) \, \mathrm{d}p = \mathrm{E}(p) \quad (!!) \end{aligned}$$

$\mathrm{E}(p)$ (and not the mode!) is the probability of every ‘future’ event which is believed to have the **same p** of the ‘previous’ ones.

(But keep in mind the **inductivist turkey**!)

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes <u>and</u> the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes <u>and</u> the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$\mathrm{E}(p) \approx p_m = \frac{x}{n}$$

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes <u>and</u> the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$
\begin{aligned}
\mathrm{E}(p) &\approx p_m = \frac{x}{n} \\
\mathrm{Var}(p) &\approx \frac{x}{n}\left(1-\frac{x}{n}\right)\frac{1}{n} = \frac{p_m\,(1-p_m)}{n}
\end{aligned}
$$

Inferring the “Bernoulli’s p”

Large number behaviour

When the number of successes and the number of failures become ‘large’ (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$
\begin{aligned}
\mathrm{E}(p) &\approx p_m = \frac{x}{n} \\
\mathrm{Var}(p) &\approx \frac{x}{n}\left(1 - \frac{x}{n}\right)\frac{1}{n} = \frac{p_m\,(1-p_m)}{n} \\
\sigma(p)(=\sigma_p) &\approx \sqrt{\frac{p_m\,(1-p_m)}{n}} \propto \frac{1}{\sqrt{n}}
\end{aligned}
$$

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes and the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$\begin{aligned}
\mathrm{E}(p) &\approx p_m = \frac{x}{n} \\
\mathrm{Var}(p) &\approx \frac{x}{n}\left(1-\frac{x}{n}\right)\frac{1}{n} = \frac{p_m\,(1-p_m)}{n} \\
\sigma(p)(=\sigma_p) &\approx \sqrt{\frac{p_m\,(1-p_m)}{n}} \propto \frac{1}{\sqrt{n}}
\end{aligned}$$

Moreover $f(p)$ tends to a Gaussian distribution:

$$p \sim \mathcal{N}(p_m, \sigma_p)$$

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes and the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$\begin{aligned}
\mathrm{E}(p) &\approx p_m = \frac{x}{n} \\
\mathrm{Var}(p) &\approx \frac{x}{n}\left(1-\frac{x}{n}\right)\frac{1}{n} = \frac{p_m\,(1-p_m)}{n} \\
\sigma(p)(=\sigma_p) &\approx \sqrt{\frac{p_m\,(1-p_m)}{n}} \propto \frac{1}{\sqrt{n}}
\end{aligned}$$

Moreover $f(p)$ tends to a Gaussian distribution:

$$p \sim \mathcal{N}(p_m, \sigma_p)$$

When $n \to \infty$, then $\sigma_p \to 0$,

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes and the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$\begin{aligned} \mathrm{E}(p) &\approx p_m = \frac{x}{n} \\ \mathrm{Var}(p) &\approx \frac{x}{n}\left(1-\frac{x}{n}\right)\frac{1}{n} = \frac{p_m\,(1-p_m)}{n} \\ \sigma(p)(=\sigma_p) &\approx \sqrt{\frac{p_m\,(1-p_m)}{n}} \propto \frac{1}{\sqrt{n}} \end{aligned}$$

Moreover $f(p)$ tends to a Gaussian distribution:

$$p \sim \mathcal{N}(p_m, \sigma_p)$$

When $n \to \infty$, then $\sigma_p \to 0$, $\to$ and hence

$$P(E_{i>n}\,|\,x, n) \text{ "}\longrightarrow\text{" } \frac{x}{n}$$

Inferring the "Bernoulli's p"

Large number behaviour

When the number of successes and the number of failures become 'large' (x large is not enough, as it can be easily understood from the simmetric properties of the binomial $p \leftrightarrow q$):

$$\begin{aligned} \mathrm{E}(p) &\approx p_m = \frac{x}{n} \\ \mathrm{Var}(p) &\approx \frac{x}{n}\left(1-\frac{x}{n}\right)\frac{1}{n} = \frac{p_m\,(1-p_m)}{n} \\ \sigma(p)(=\sigma_p) &\approx \sqrt{\frac{p_m\,(1-p_m)}{n}} \propto \frac{1}{\sqrt{n}} \end{aligned}$$

Moreover $f(p)$ tends to a Gaussian distribution:

$$p \sim \mathcal{N}(p_m, \sigma_p)$$

When $n \to \infty$, then $\sigma_p \to 0$, $\to$ and hence

$$P(E_{i>n}\,|\,x, n) \text{ "}\longrightarrow\text{" } \frac{x}{n}$$

(Similarly to Bernoulli's theorem, it is not a 'mathematical' limit!)

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;
- and $(n - x)$ large

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;
- and $(n - x)$ large
  (remember: in the binomial what is 'success' and what is 'failure' is not absolute: $p \longleftrightarrow q = 1 - p$)

Inferring the “Bernoulli's p”

Large number behaviour: summary

When

- n large;
- x large;
- <u>and</u> $(n - x)$ large
  (remember: in the binomial what is ‘success’ and what is ‘failure’
  is not absolute: $p \longleftrightarrow q = 1 - p$),

then

$$\mathrm{E}(p) \approx \frac{x}{n}$$

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;
- and $(n - x)$ large
  (remember: in the binomial what is 'success' and what is 'failure' is not absolute: $p \longleftrightarrow q = 1 - p$),

then

$$E(p) \approx \frac{x}{n}$$

$$\sigma(p) \approx \frac{1}{\sqrt{n}}\sqrt{\frac{x}{n}\left(1 - \frac{x}{n}\right)}$$

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;
- and $(n-x)$ large
  (remember: in the binomial what is 'success' and what is 'failure'
  is not absolute: $p \longleftrightarrow q = 1-p$),

then

$$\begin{aligned} \mathrm{E}(p) &\approx \frac{x}{n} \\ \sigma(p) &\approx \frac{1}{\sqrt{n}}\sqrt{\frac{x}{n}\left(1-\frac{x}{n}\right)} \end{aligned}$$

— $f(p\,|\,x,n)$ tends to Gaussian

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;
- and $(n-x)$ large
  (remember: in the binomial what is 'success' and what is 'failure' is not absolute: $p \longleftrightarrow q = 1-p$),

then

$$\begin{aligned} \mathrm{E}(p) &\approx \frac{x}{n} \\ \sigma(p) &\approx \frac{1}{\sqrt{n}}\sqrt{\frac{x}{n}\left(1-\frac{x}{n}\right)} \end{aligned}$$

— $f(p\,|\,x,n)$ tends to Gaussian,
a reflection of the Gaussian limit of $f(x\,|\,p,n)$

Inferring the "Bernoulli's p"

Large number behaviour: summary

When

- n large;
- x large;
- and $(n-x)$ large
  (remember: in the binomial what is 'success' and what is 'failure' is not absolute: $p \longleftrightarrow q = 1-p$),

then

$$\mathrm{E}(p) \approx \frac{x}{n}$$
$$\sigma(p) \approx \frac{1}{\sqrt{n}}\sqrt{\frac{x}{n}\left(1-\frac{x}{n}\right)}$$

— $f(p\,|\,x, n)$ tends to Gaussian,
a reflection of the Gaussian limit of $f(x\,|\,p, n)$

— **The probability of a future events is evaluated from the relative frequency of the past events**

Inferring the “Bernoulli’s p”

Large number behaviour: summary

When

- n large;
- x large;
- and $(n - x)$ large
  (remember: in the binomial what is ‘success’ and what is ‘failure’ is not absolute: $p \longleftrightarrow q = 1 - p$),

then

$$\begin{aligned} \mathrm{E}(p) &\approx \frac{x}{n} \\ \sigma(p) &\approx \frac{1}{\sqrt{n}}\sqrt{\frac{x}{n}\left(1 - \frac{x}{n}\right)} \end{aligned}$$

— $f(p \,|\, x, n)$ tends to Gaussian,
a reflection of the Gaussian limit of $f(x \,|\, p, n)$

— **The probability of a future events is evaluated from the relative frequency of the past events**

— **No need of ‘frequentistic definition’!**

Probability vs frequency

Frequency and probability are **related** in probability theory:

Probability vs frequency

Frequency and probability are **related** in probability theory:

- Relative frequencies of successes in future trials can be 'forecasted' from p

Probability vs frequency

Frequency and probability are **related** in probability theory:

- Relative frequencies of successes in future trials can be 'forecasted' from p (Bernoulli theorem).

Probability vs frequency

Frequency and probability are **related** in probability theory:

- Relative frequencies of successes in future trials can be ‘forecasted’ from p (Bernoulli theorem).
- Probability p can be evaluated from past frequencies, under some assumptions

Probability vs frequency

Frequency and probability are **related** in probability theory:

- Relative frequencies of successes in future trials can be 'forecasted' from p (Bernoulli theorem).
- Probability p can be evaluated from past frequencies, under some assumptions ('Bayes theorem' $\rightarrow$ **Laplace's rule**)

Probability vs frequency

Frequency and probability are **related** in probability theory:

- Relative frequencies of successes in future trials can be 'forecasted' from p (Bernoulli theorem).
- Probability p can be evaluated from past frequencies, under some assumptions ('Bayes theorem' $\rightarrow$ **Laplace's rule**)

BUT

- There is no need to identify the two concepts.

Probability vs frequency

Frequency and probability are **related** in probability theory:

- Relative frequencies of successes in future trials can be 'forecasted' from p (Bernoulli theorem).
- Probability p can be evaluated from past frequencies, under some assumptions ('Bayes theorem' $\rightarrow$ **Laplace's rule**)

BUT

- There is no need to identify the two concepts.
- It does not justify the frequentistic definition.