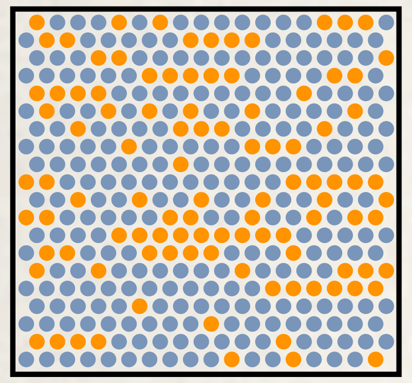
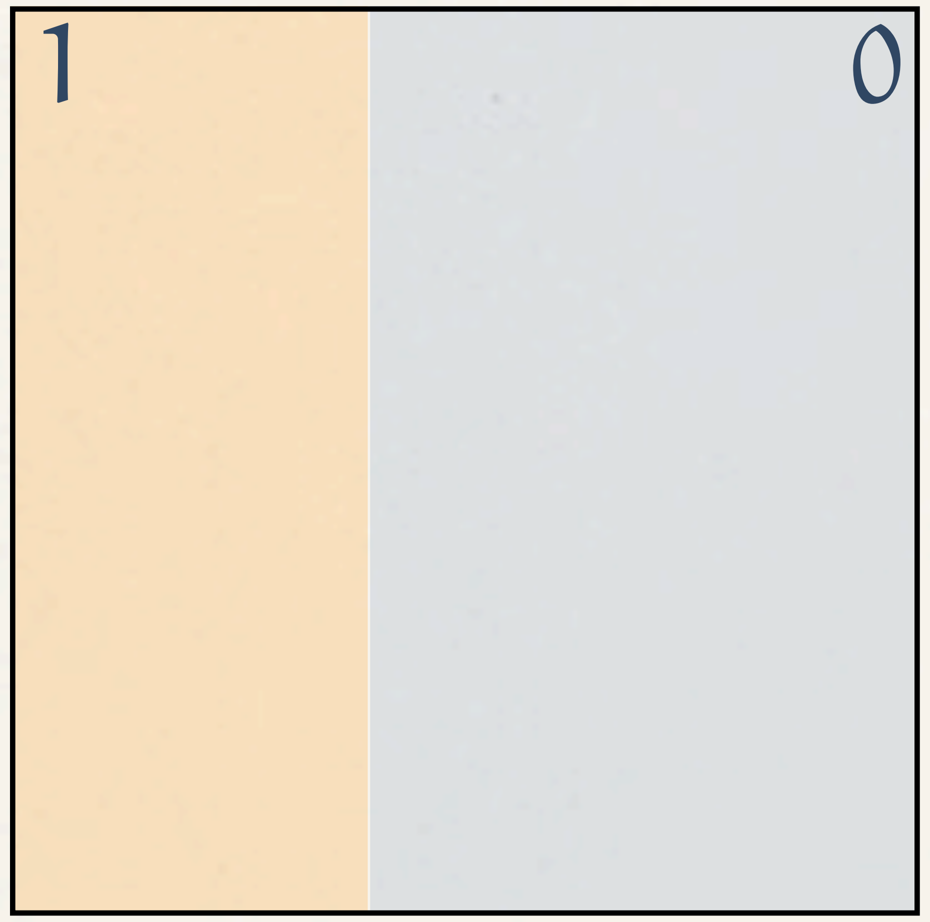
Tableau 11, Part 1

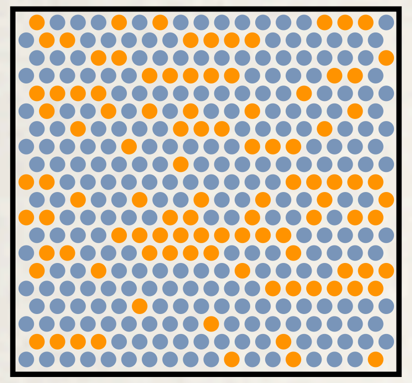
The fabulous limit laws
Chebyshev's enduring inequality, the magisterial law of large numbers



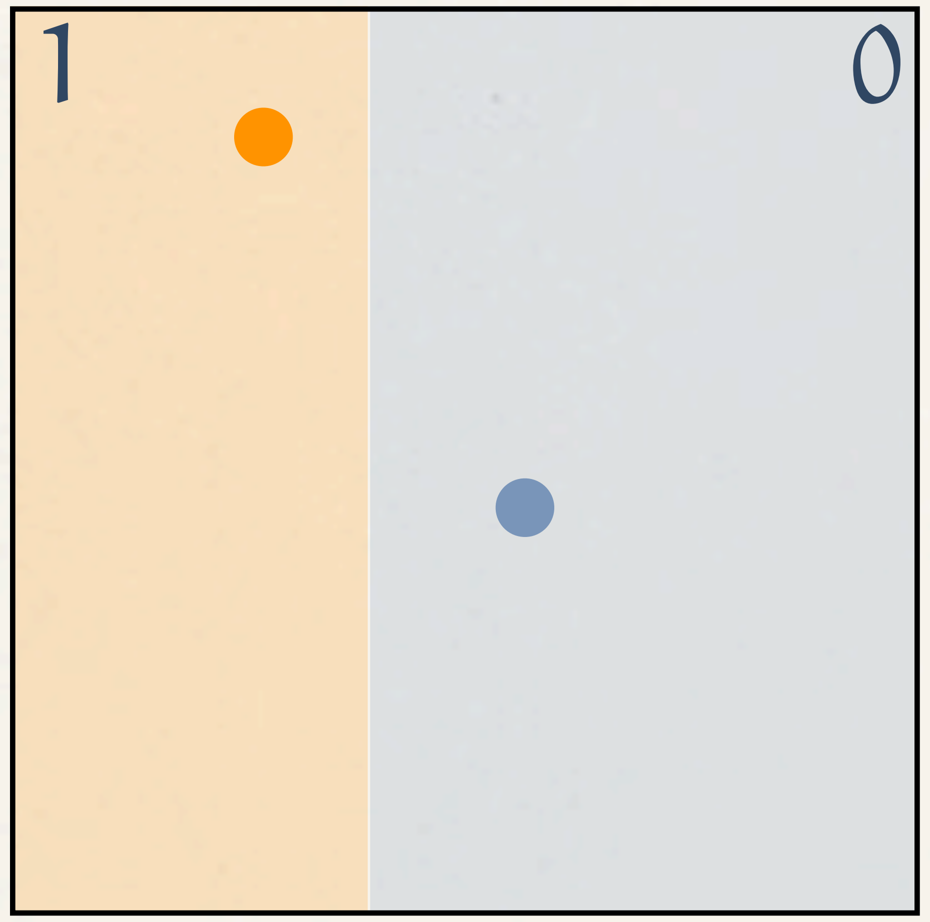
Fixed but unknown proportion p



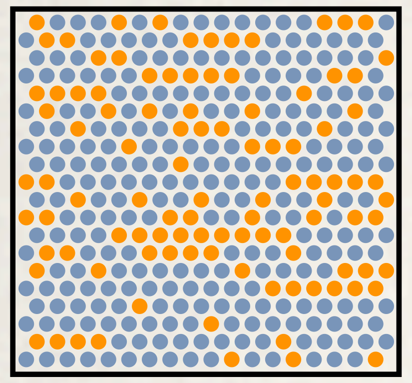
Fixed but unknown proportion $q = 1 - p$



Fixed but unknown proportion p

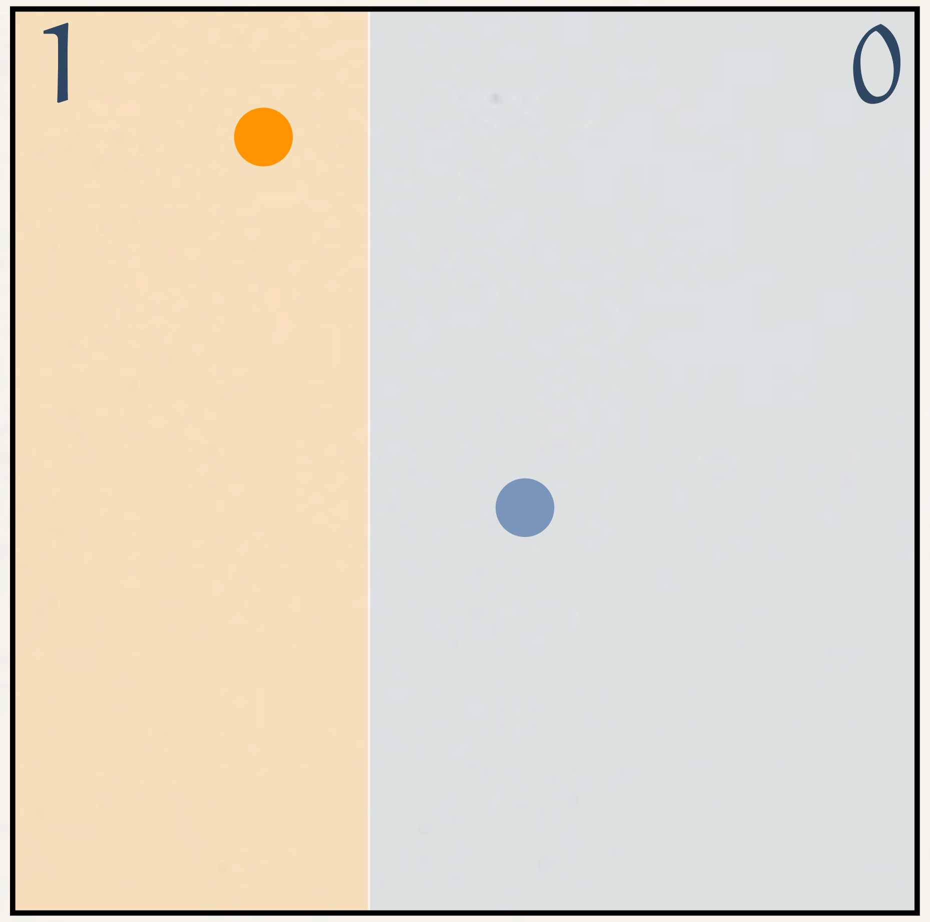


Fixed but unknown proportion q = 1 – p

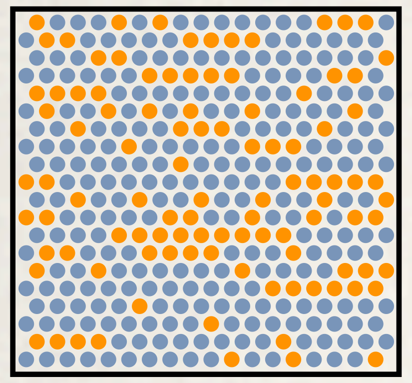


Sampling from a dichotomous population

Fixed but unknown proportion p

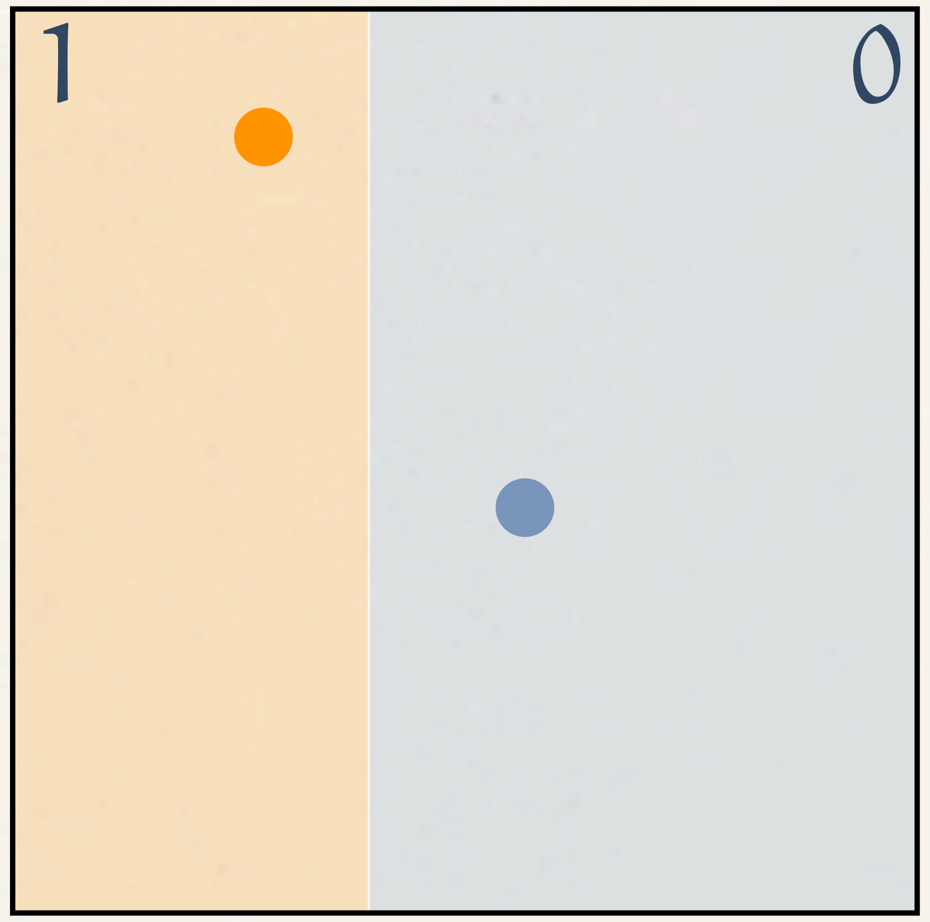


Fixed but unknown proportion $q = 1 - p$



Sampling from a dichotomous population

Fixed but unknown proportion p



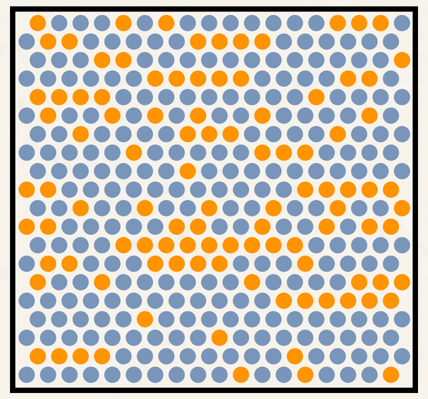
Fixed but unknown proportion $q = 1 - p$

Bernoulli trial with success probability p

$$X \sim \text{Bernoulli}(p)$$

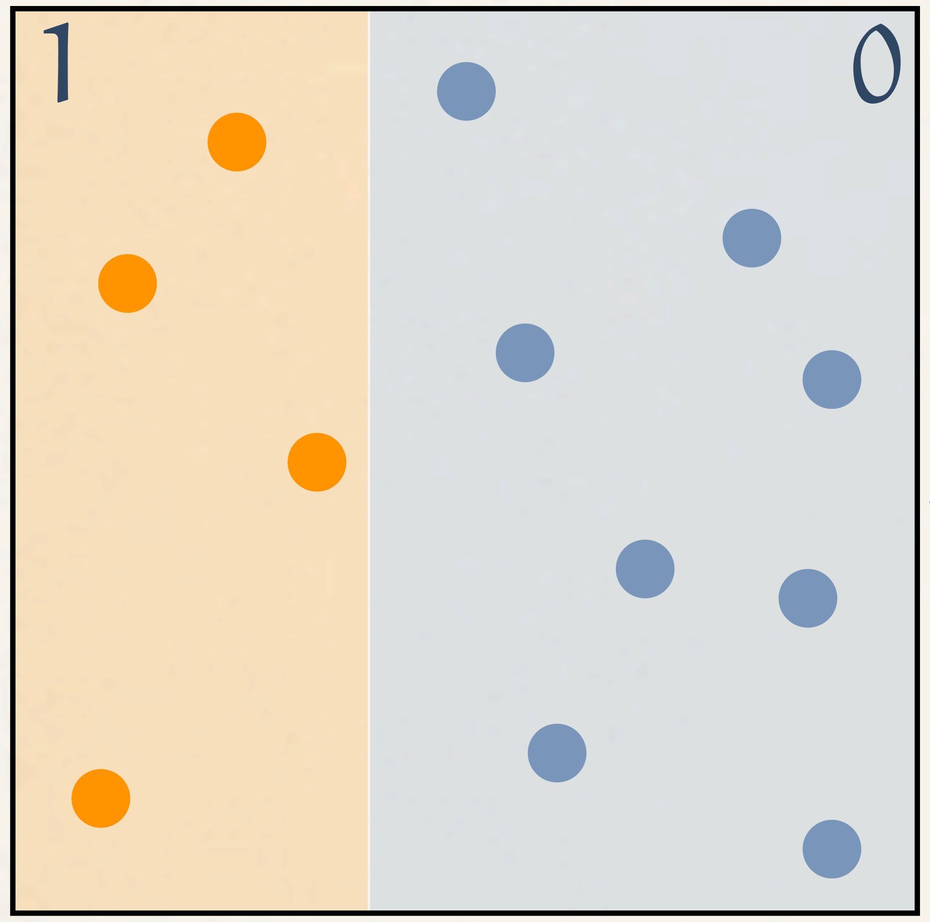
$$X = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } q. \end{cases}$$

A model for a poll



Random sample: repeated independent trials

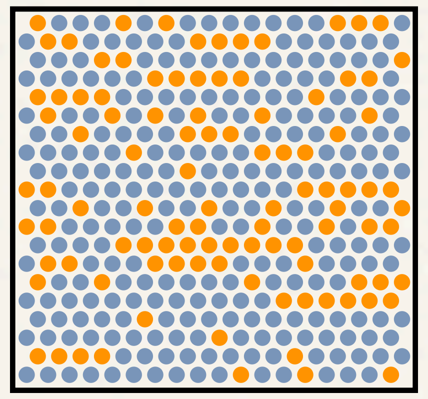
Fixed but unknown proportion p



Fixed but unknown proportion q = 1 − p

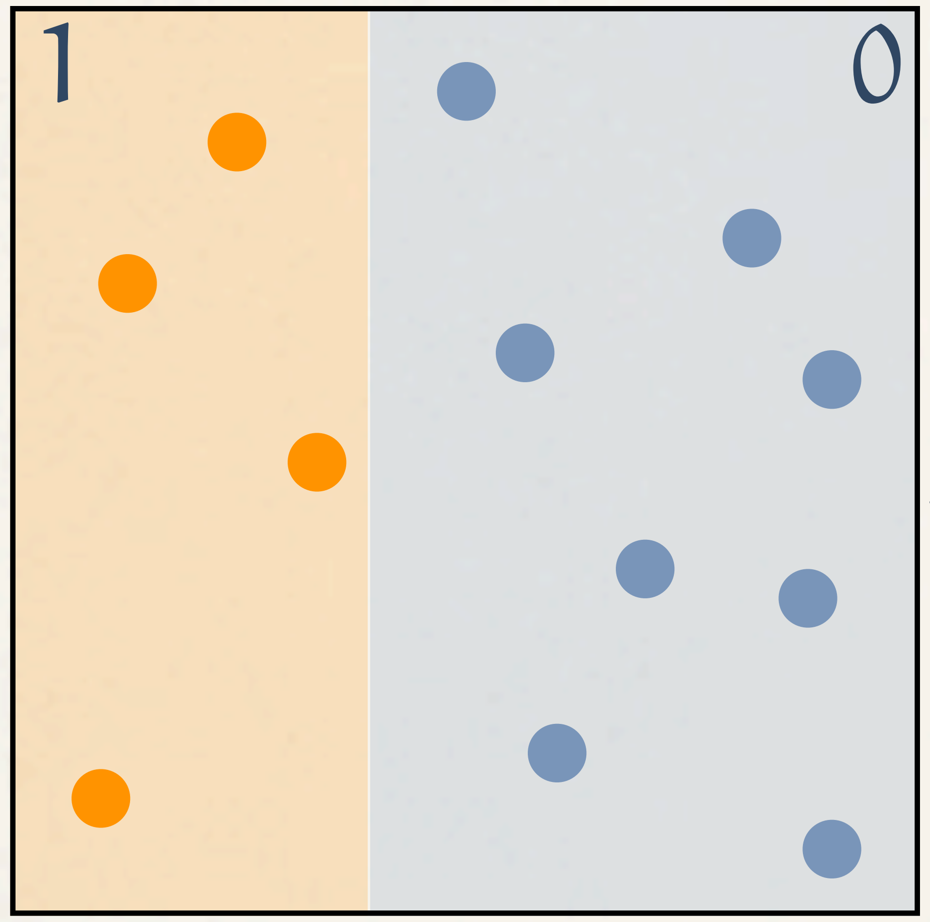
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
0	1	1	0	0	0	0	0	0	0	1	0

A model for a poll



Random sample: repeated independent trials

Fixed but unknown proportion p

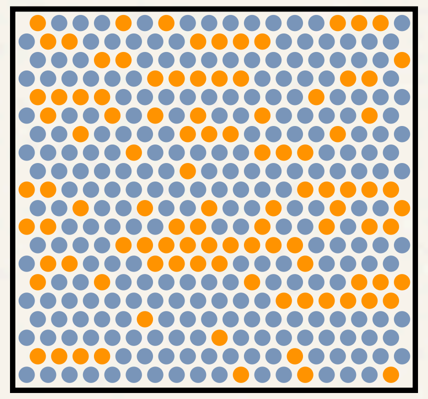


Fixed but unknown proportion $q = 1 - p$

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
0	1	1	0	0	0	0	0	0	0	1	0

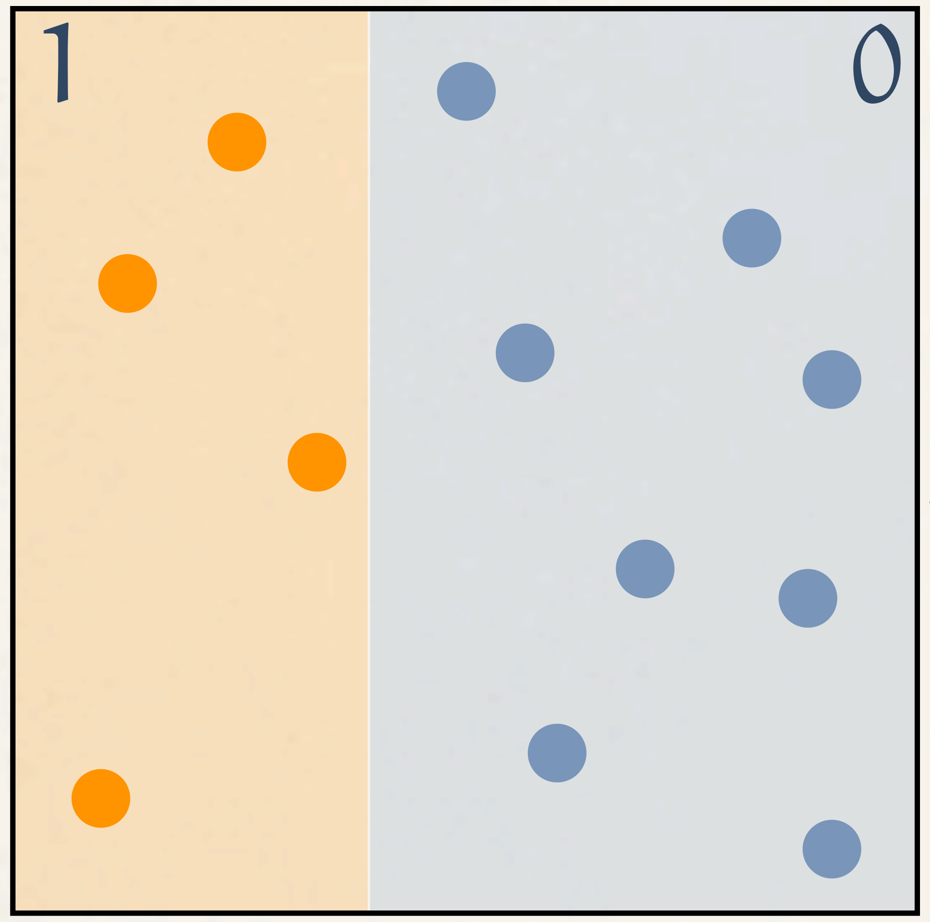
$$\text{Bernoulli}(p)\text{ trials:}\quad X_1, X_2, \dots, X_n = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } q. \end{cases}$$

A model for a poll



Random sample: repeated independent trials

Fixed but unknown proportion p



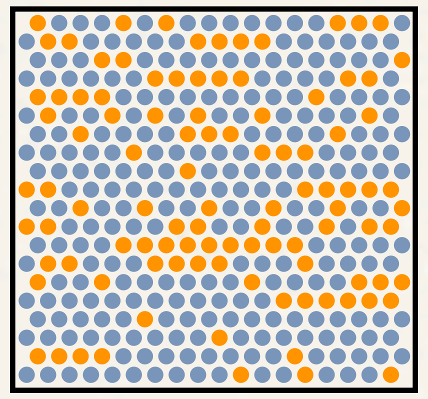
Fixed but unknown proportion $q = 1 - p$

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
0	1	1	0	0	0	0	0	0	0	1	0

Bernoulli(p) trials: $X_1, X_2, \dots, X_n = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } q. \end{cases}$

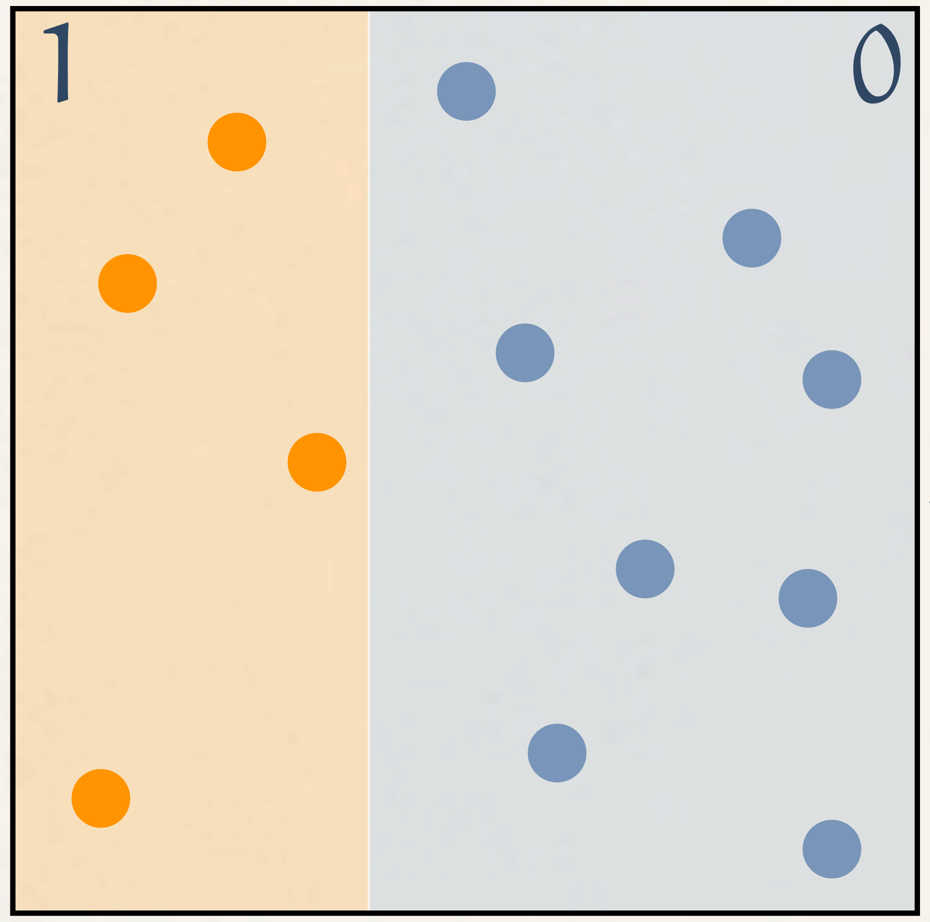
Accumulated successes: $S_n = X_1 + X_2 + \cdots + X_n$

A model for a poll



Random sample: repeated independent trials

Fixed but unknown proportion p



Fixed but unknown proportion $q = 1 - p$

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
0	1	1	0	0	0	0	0	0	0	1	0

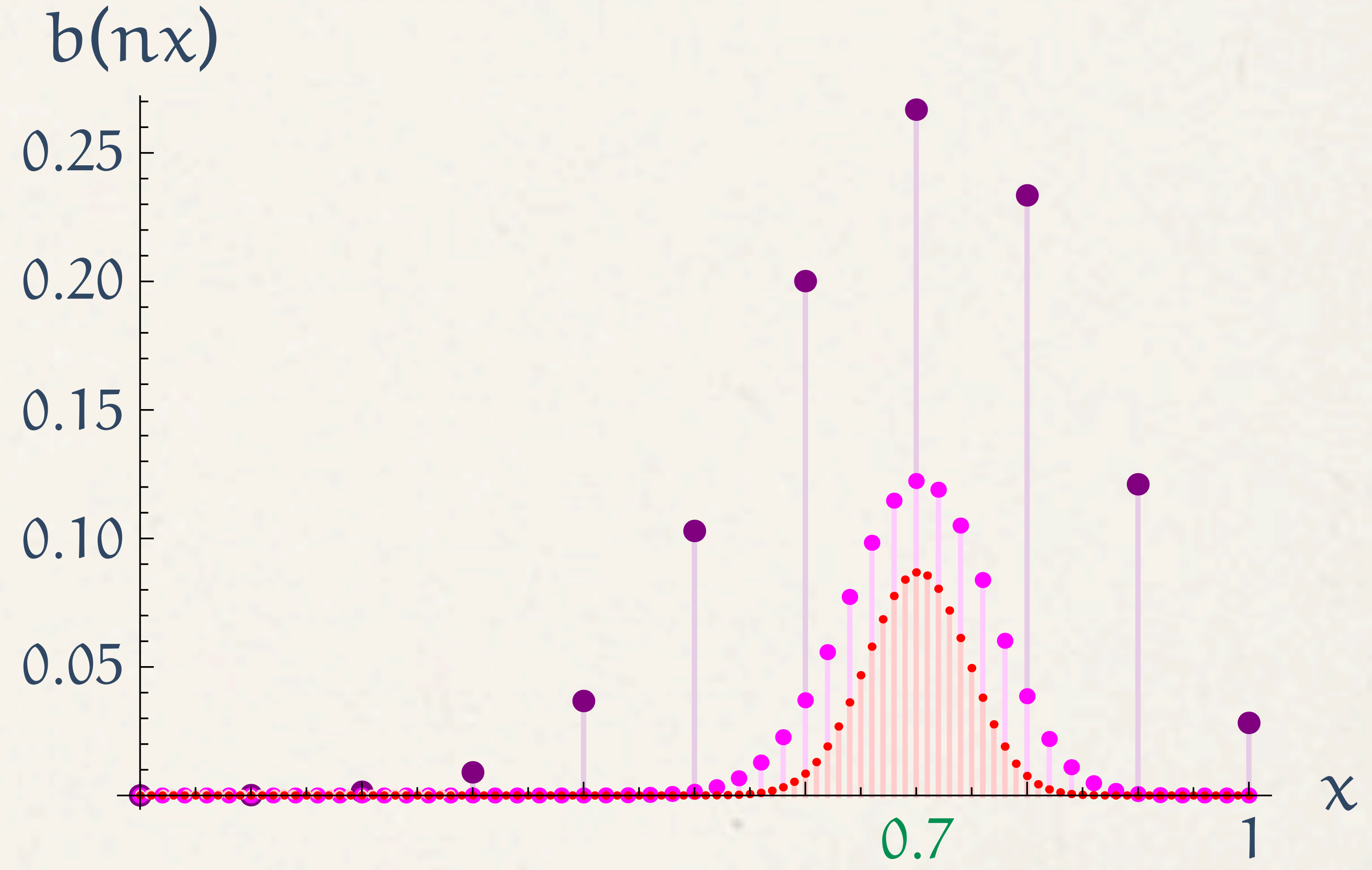
Bernoulli(p) trials: $X_1, X_2, \dots, X_n = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } q. \end{cases}$

Accumulated successes: $S_n = X_1 + X_2 + \cdots + X_n$

Is the relative frequency of accumulated successes in the sample, S_n/n, a good approximation to the fixed but unknown population proportion p?

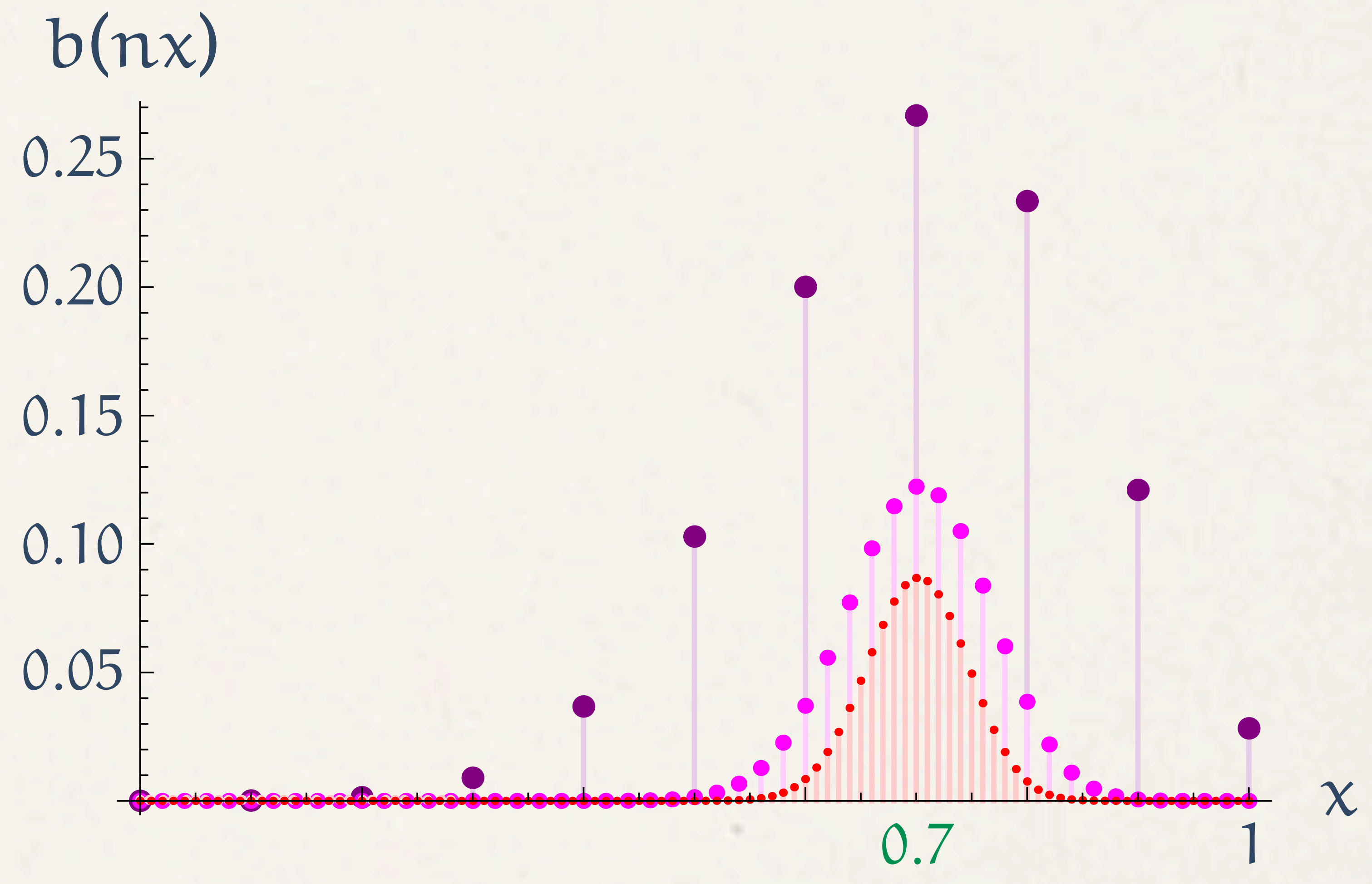
Reasons for optimism

Reasons for optimism



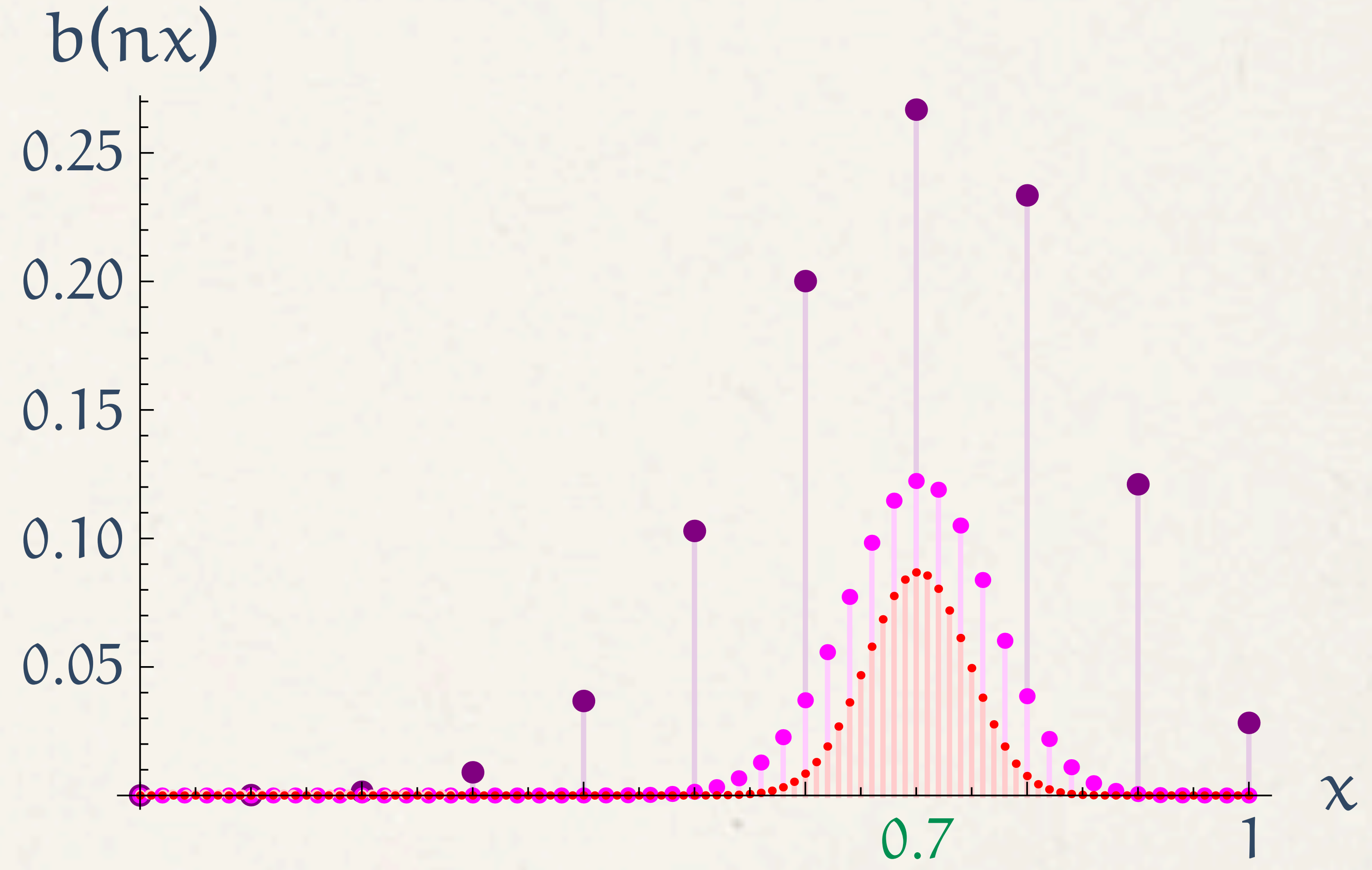
Reasons for optimism

- Maximum likelihood: The value S_n/n is the maximum likelihood estimate of the bias p: it gives the largest a posteriori probability of obtaining the observed number of successes S_n.



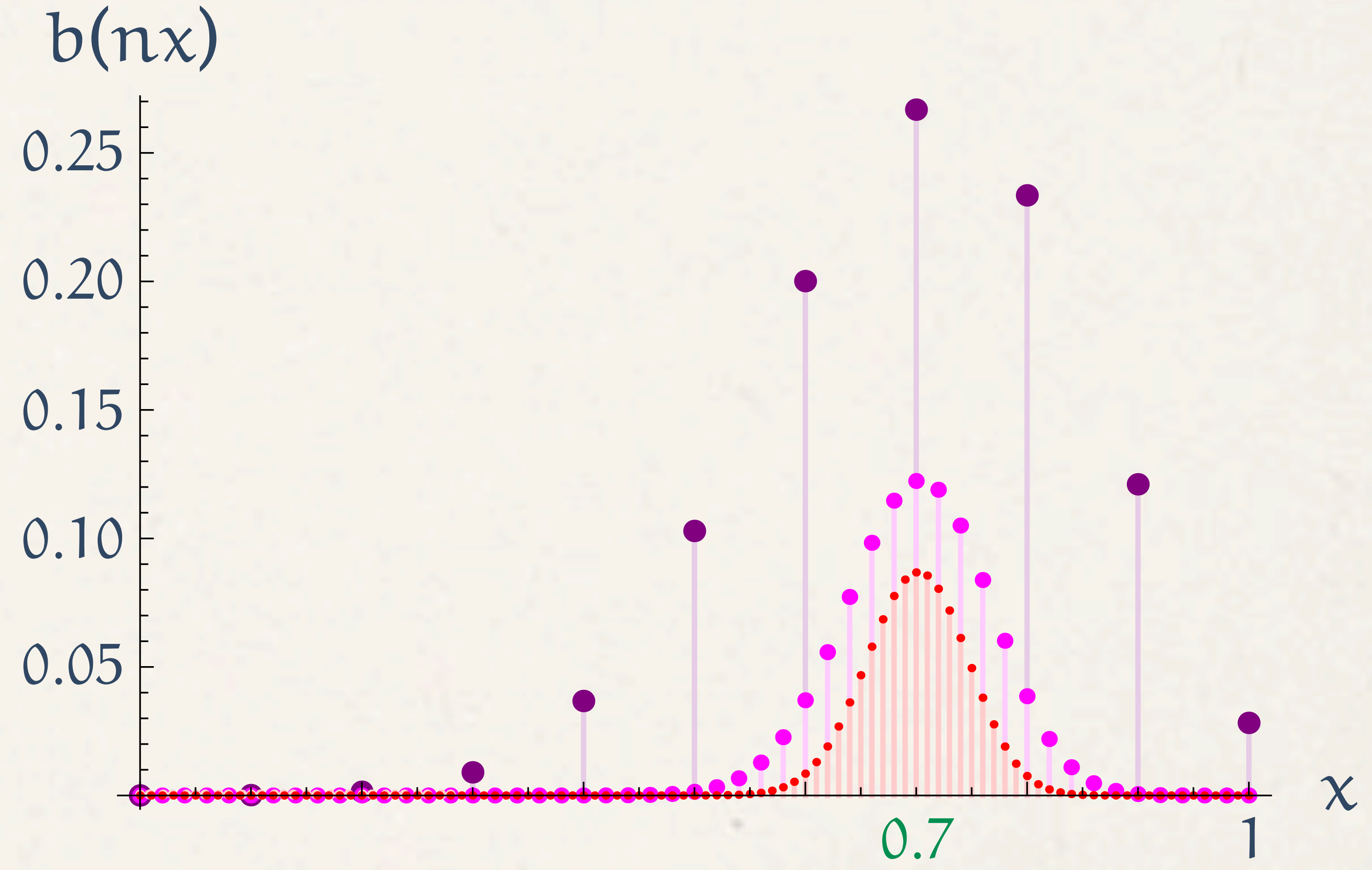
Reasons for optimism

- Maximum likelihood: The value S_n/n is the maximum likelihood estimate of the bias p: it gives the largest a posteriori probability of obtaining the observed number of successes S_n.
- Maxima: The mass function of S_n/n attains its maximum value in a small neighbourhood of p.



Reasons for optimism

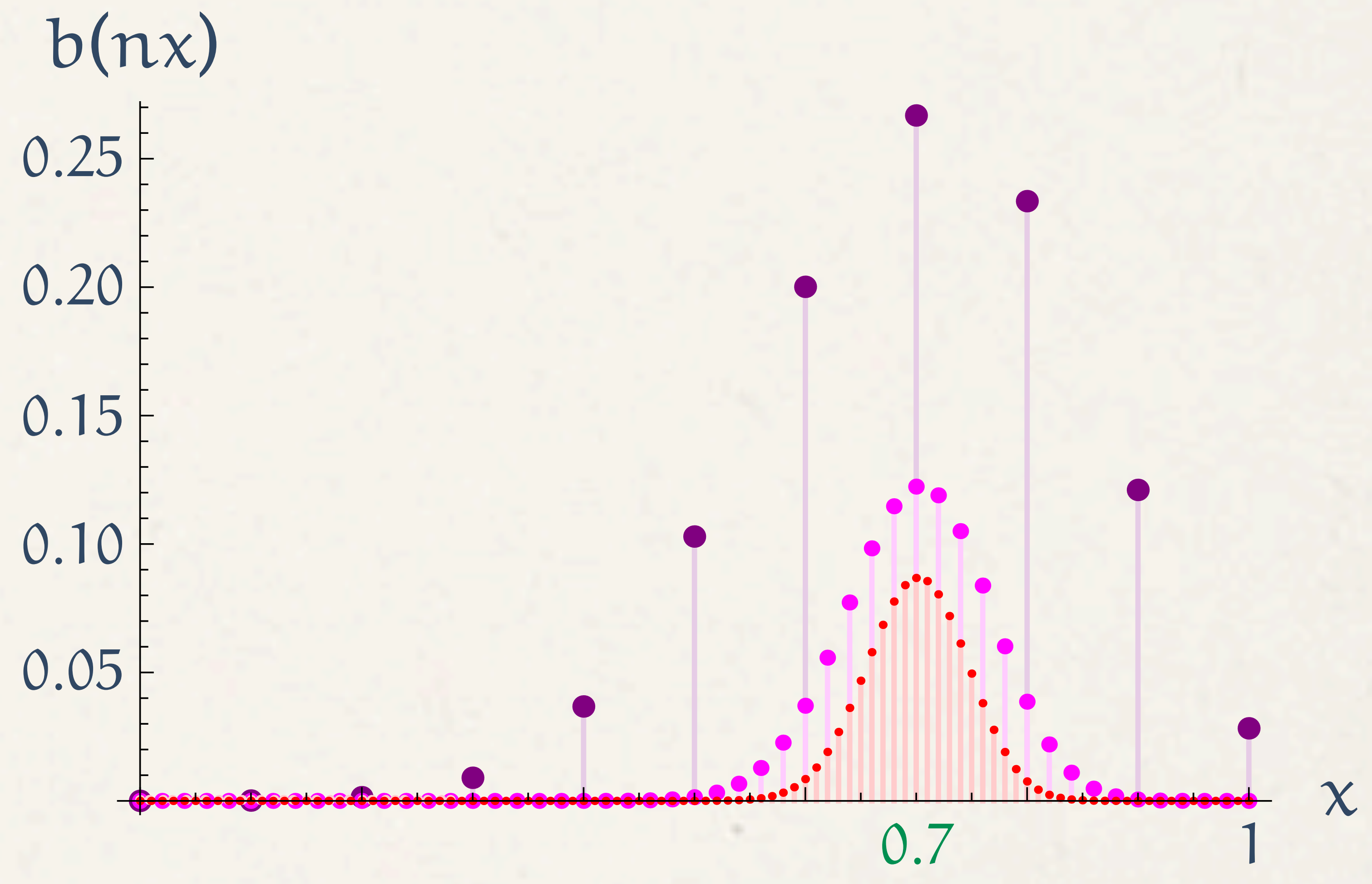
- Maximum likelihood: The value S_n/n is the maximum likelihood estimate of the bias p: it gives the largest a posteriori probability of obtaining the observed number of successes S_n.
- Maxima: The mass function of S_n/n attains its maximum value in a small neighbourhood of p.
- Expectation: Its expectation is p: $\mathbf{E}(S_n/n) = p$.



Reasons for optimism

- Maximum likelihood: The value S_n/n is the maximum likelihood estimate of the bias p: it gives the largest a posteriori probability of obtaining the observed number of successes S_n.
- Maxima: The mass function of S_n/n attains its maximum value in a small neighbourhood of p.
- Expectation: Its expectation is p: $\mathbf{E}(S_n/n) = p$.
- Variance: Graphical evidence suggests that the mass function of S_n/n gets increasingly concentrated in a small neighbourhood of p when n is large. The variance provides more evidence in support of this observation:

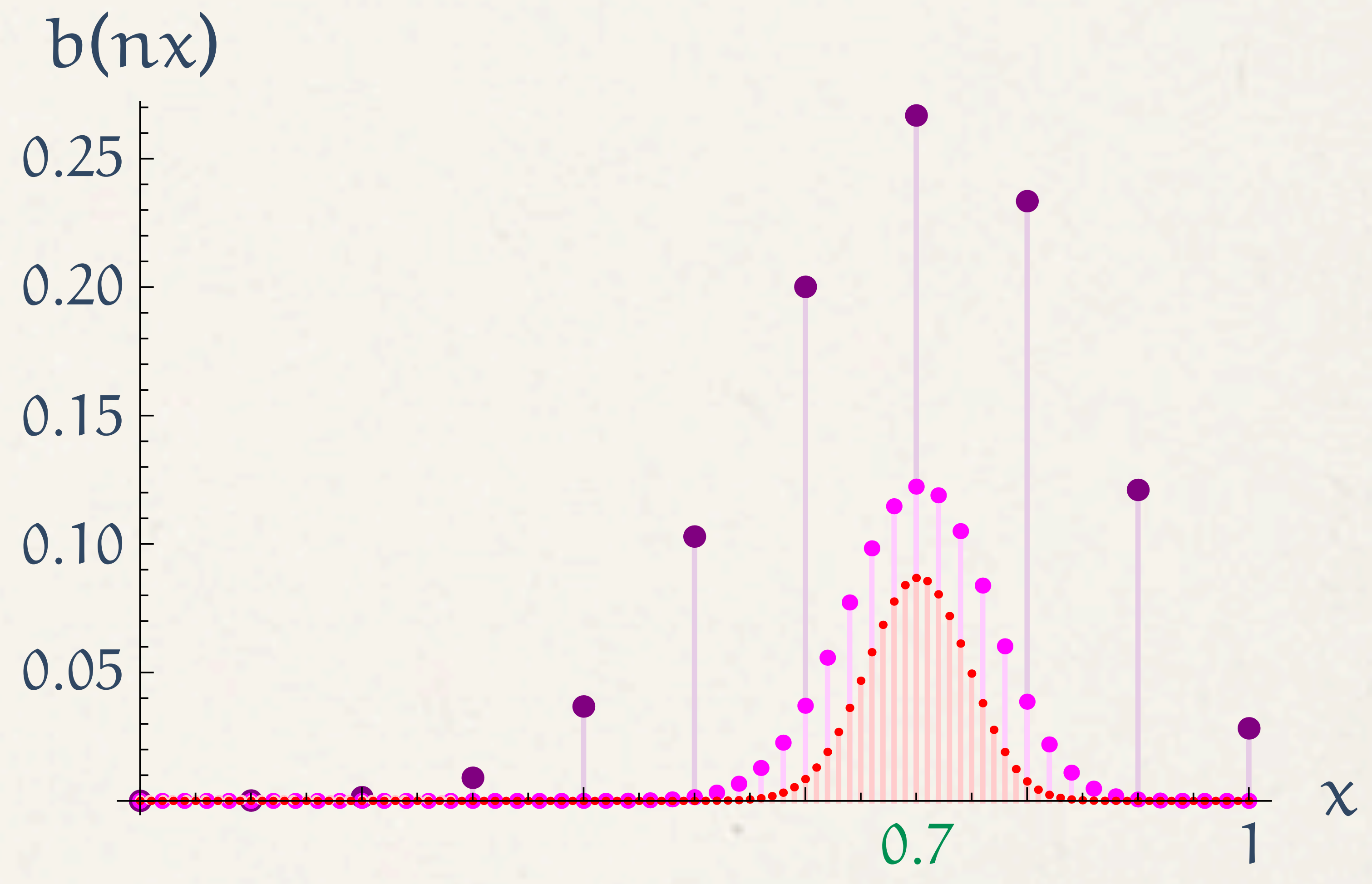
$$\mathrm{Var}(S_n) = npq \quad \text{or} \quad \frac{1}{n}\sqrt{\mathrm{Var}(S_n)} = \frac{\sqrt{pq}}{\sqrt{n}}.$$



Reasons for optimism

- Maximum likelihood: The value S_n/n is the maximum likelihood estimate of the bias p: it gives the largest a posteriori probability of obtaining the observed number of successes S_n.
- Maxima: The mass function of S_n/n attains its maximum value in a small neighbourhood of p.
- Expectation: Its expectation is p: $\mathbf{E}(S_n/n) = p$.
- Variance: Graphical evidence suggests that the mass function of S_n/n gets increasingly concentrated in a small neighbourhood of p when n is large. The variance provides more evidence in support of this observation:

$$\operatorname{Var}(S_n) = npq \quad \text{or} \quad \frac{1}{n}\sqrt{\operatorname{Var}(S_n)} = \frac{\sqrt{pq}}{\sqrt{n}}.$$



But how good is the approximation S_n/n really?