# intro to data science with probability & statistics

# CSCI 3022

Lecture 8
February 7, 2017

More Discrete RVs!
- Binomial
- Geometric
- Negative Binomial
- Poisson

Department of Computer Science
UNIVERSITY OF COLORADO BOULDER

Dan Larremore

# Last time on CSCI 3022:

- A discrete RV $X$ has a **Bernoulli distribution** with parameter $p$ ($0 \le p \le 1$) if its probability mass function (PMF) is given by:

$$p_X(1) = P(X = 1) = p \qquad p_X(0) = P(X = 0) = 1 - p$$

$f(1) = p$ $\qquad$ $f(0) = 1 - p$

# The Binomial Distribution

- Suppose I sum 5 Bernoulli random variables. This sum is a RV too!
- It takes on values in the interval [0,5].
- What is the PMF of the sum of 5 Bernoulli RVs? Let's build it up!

$f(0) = P(\text{all are T}) = (1-p)^5$

$f(1) = P(1H, 4T) = 5p^1(1-p)^4$

$f(2) = P(2H, 3T) = \binom{5}{2} p^2 (1-p)^3$

$$\left.\begin{matrix} H & T & T & T & T \\ T & H & T & T & T \\ T & T & H & T & T \\ T & T & T & H & T \\ T & T & T & T & H \end{matrix}\right] \rightarrow \binom{5}{1}$$

# The Binomial Distribution

- The sum of Bernoulli RVs is the **Binomial Distribution** (see top of slide).
- It is the distribution of the number of "heads" you'll get when flipping a coin n times.
- Note that it is parameterized by the number of flips n and the Bernoulli parameter p. So we call this *Bin(n,p)*.

k heads

n-k tails

$$P(X=k;n,p)=\boxed{\binom{n}{k}p^{k}(1-p)^{n-k}}$$

# ways to have k H in n flips

# The Binomial Distribution

- The sum of Bernoulli RVs is the **Binomial Distribution** (see top of slide).
- It is the distribution of the number of “heads” you’ll get when flipping a coin n times.
- Note that it is parameterized by the number of flips n and the Bernoulli parameter p. So we call this *Bin(n,p)*.

$$f(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- **Practice**: what’s the probability that a biased coin with $\underbrace{p{=}0.8}_{p}$ comes up heads $\underbrace{7}_{k}$ times in $\underbrace{10 \text{ flips}}_{n}$?

$$= \binom{10}{7} 0.8^7 (0.2)^3$$

# The Binomial Distribution

$p = \frac{1}{8}$

- **Practice**: You are the caretaker of 50 otters at the aquarium. 100% of the otters are magnificent and intelligent. They *get* you. And you get them. It's a great job, and you feel lucky to have it. However, each otter has a 12.5% probability of being grumpy on any given day. Today, you choose 20 otters at random to put in the Display Habitat.

- What is the probability that 5 of the Display Habitat otters will be grumpy?

  $n = 20,\ k = 5,\ p = 1/8$ $\binom{20}{5}\left(\frac{1}{8}\right)^5\left(\frac{7}{8}\right)^{15}$

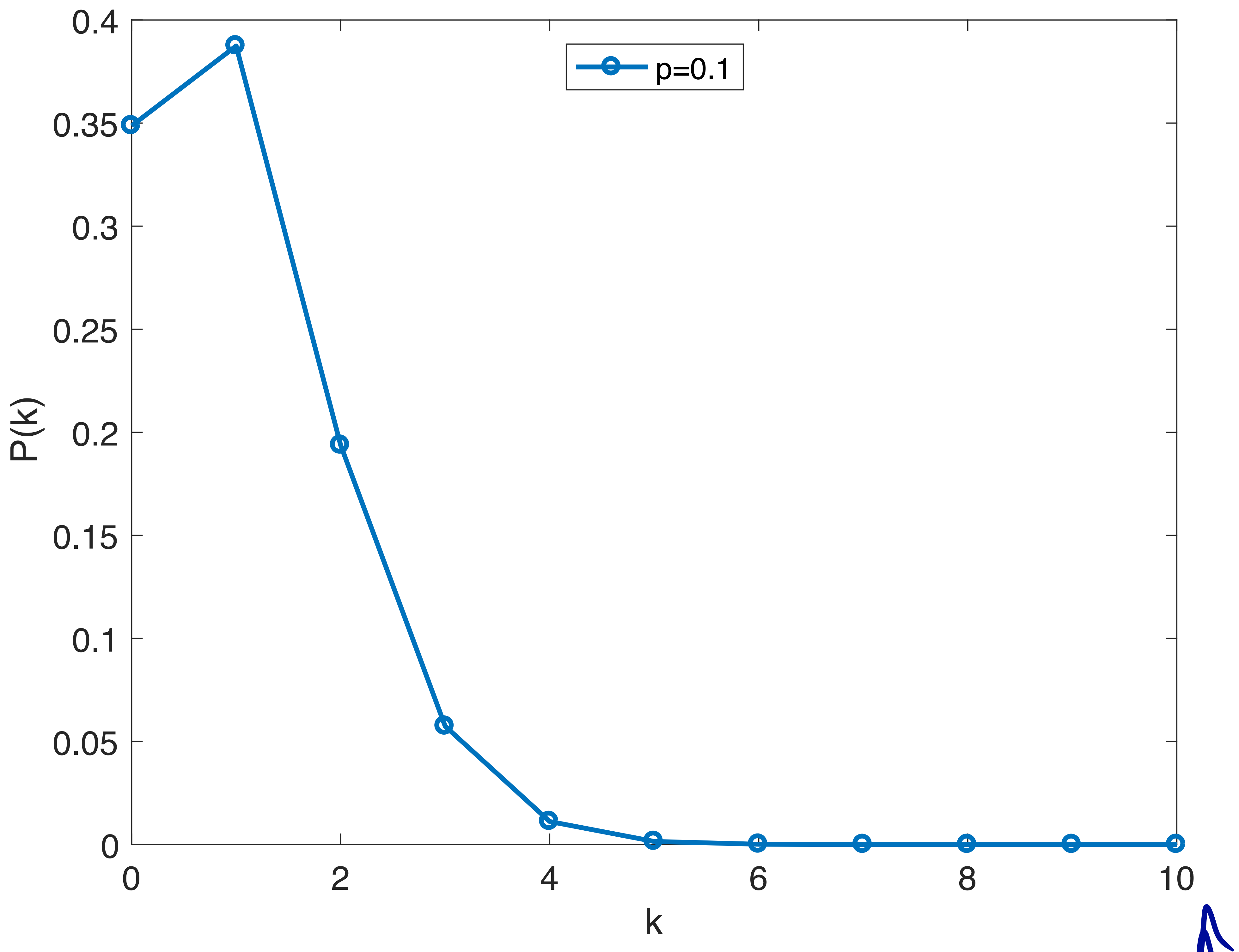
- What is the probability that all the Display Habitat otters will not be grumpy?

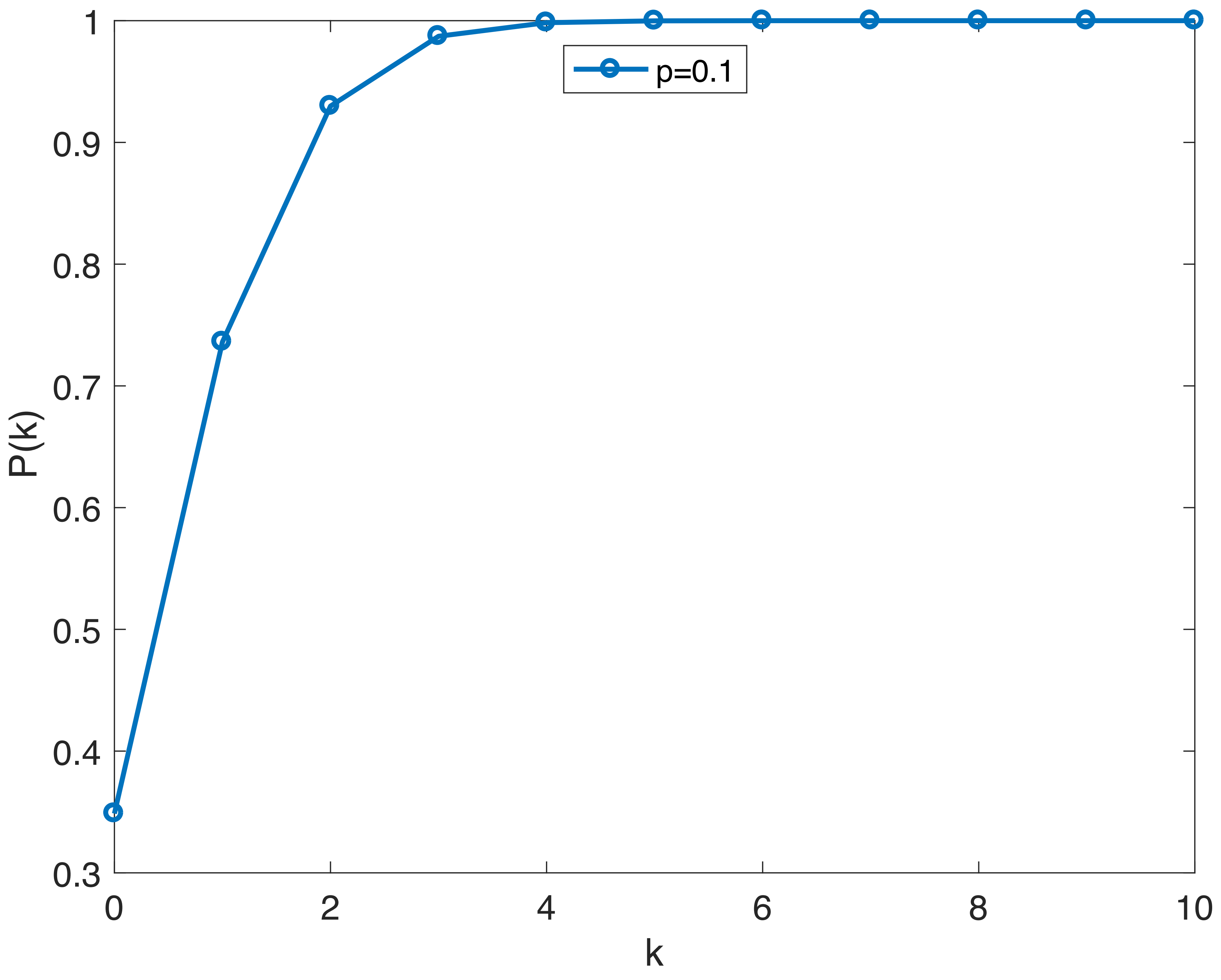
  $\binom{20}{0}\left(\frac{1}{8}\right)^0(7/8)^{20} = (7/8)^{20}$

- What are we assuming when we answer these Qs using the binomial distr.?

  Independence of otters' moods.

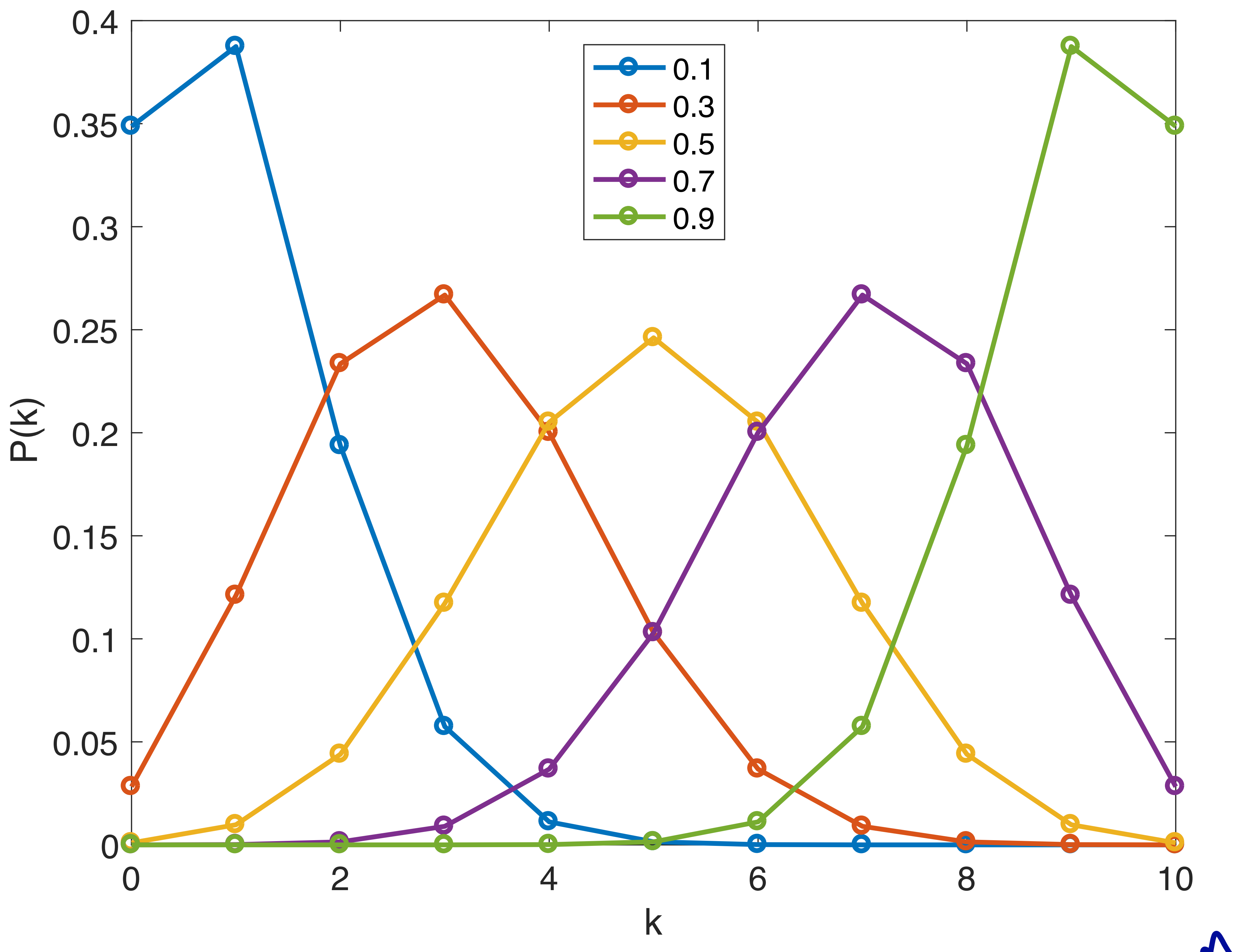
# Properties of the Binomial PMF

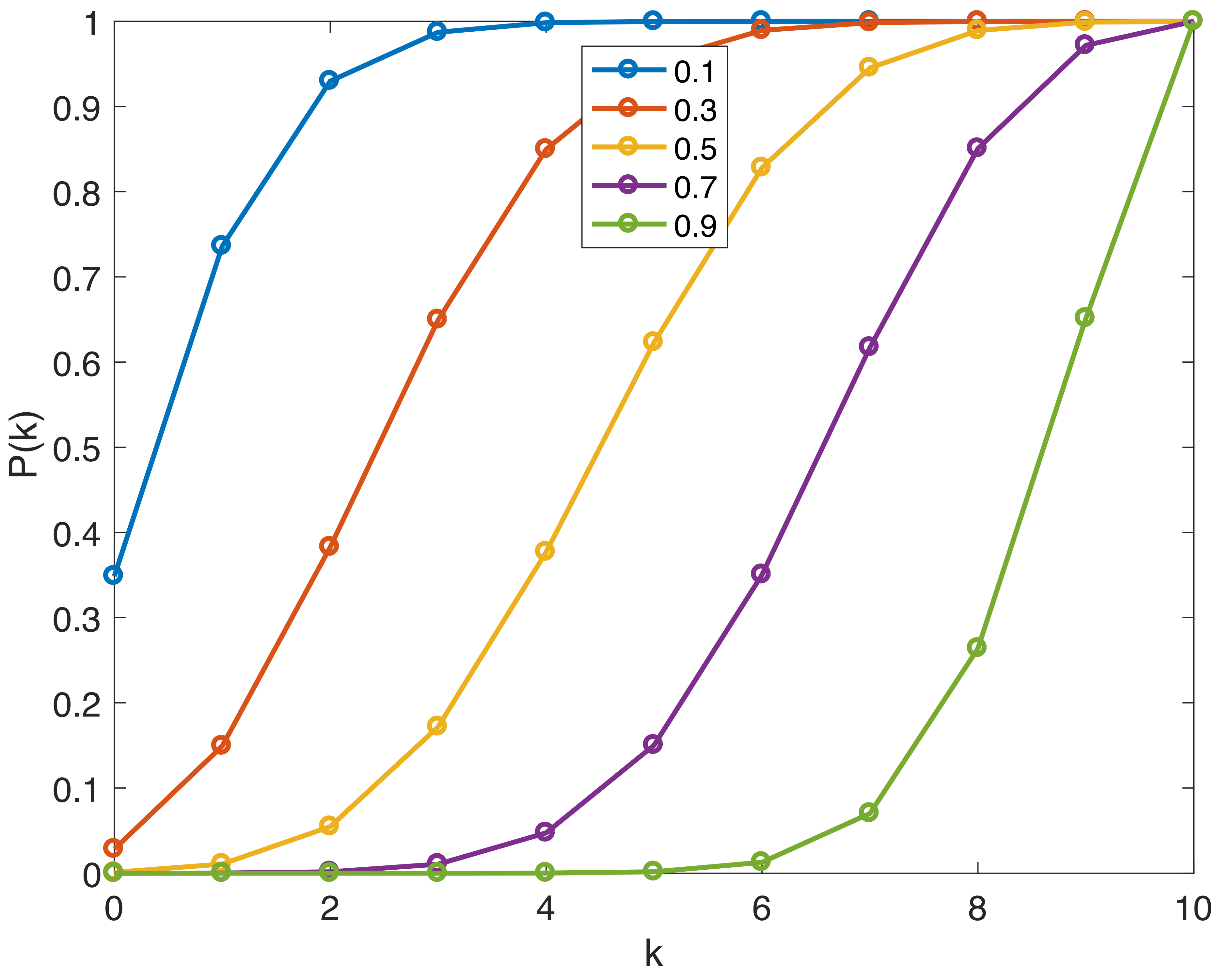



n=10

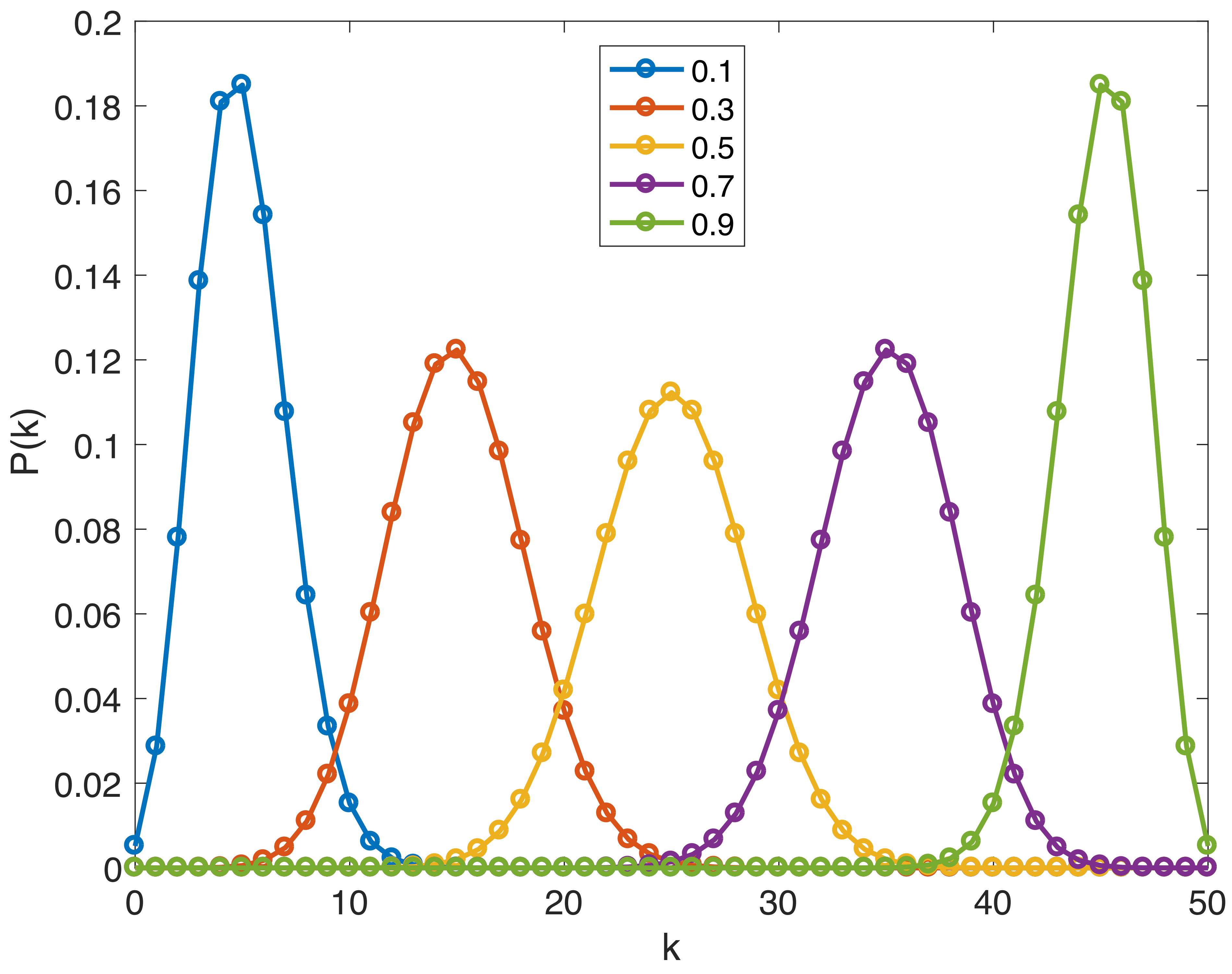
# Properties of the Binomial PMF

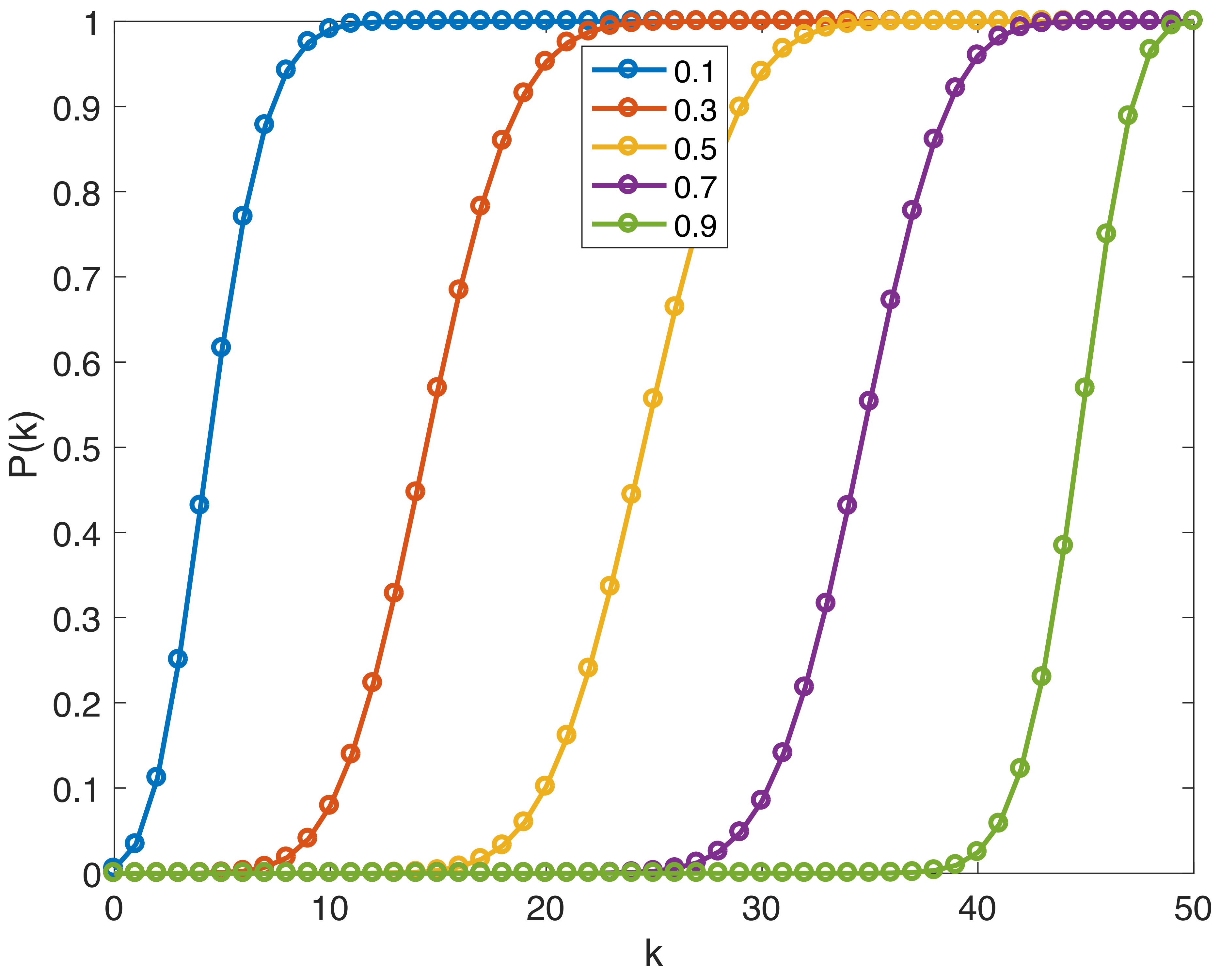



n=10

# Properties of the Binomial PMF




n=50

# By the spirit of the Binomial…

- **Today**: Let's get some more mileage out of the idea of coin flips! In the same spirit as the Binomial, today we're going to talk about:
    - **Geometric distribution** (how many flips till I get a heads?)
    - **Negative binomial distribution** (how many flips till I get 3 heads?)
    - **Poisson distribution** (how many texts do I get during class?)

# Election Day Exit Polls

- **Example**: You are doing an exit poll outside of a voting station on Election Day. As people exit, you ask them questions about their political affiliation, who they voted for, etc. In particular, you're interested in how registered Independents voted. Being well-prepared, you know that about 20% of registered voters are registered as Independents.

- **Goal 1**: Suppose you hang out until you interview 100 people. Let X be a random variable describing the number of actual Independents that you encounter.

$$X \sim \text{Bin}(0.2, 100)$$

# Election Day Exit Polls

- **Example**: You are doing an exit poll outside of a voting station on Election Day. As people exit, you ask them questions about their political affiliation, who they voted for, etc. In particular, you're interested in how registered Independents voted. Being well-prepared, you know that about 20% of registered voters are registered as Independents.

- **Goal 2**: Suppose you're talking to a lot of registered Republicans and Democrats, but haven't met an Independent yet. Let X be a random variable describing the number of people you have interviewed up to and including your first registered Independent voter.

$X \sim geo(0.2)$

# Election Day Exit Polls

- **Example**: You are doing an exit poll outside of a voting station on Election Day. As people exit, you ask them questions about their political affiliation, who they voted for, etc. In particular, you're interested in how registered Independents voted. Being well-prepared, you know that about 20% of registered voters are registered as Independents.

- **Goal 3**: You're really interested in talking to a lot of Independents. Let X be the random variable describing the number of people you have to talk to in order to interview exactly 100 registered independents.

$X \sim NB(0.2, 100)$

# Election Day Exit Polls

- **Example**: You are doing an exit poll outside of a voting station on Election Day. As people exit, you ask them questions about their political affiliation, who they voted for, etc. In particular, you're interested in how registered Independents voted. Being well-prepared, you know that about 20% of registered voters are registered as Independents.

- **Goal 4**: You're concerned about being overwhelmed during the peak voting time, so you keep an eye on the number of people arriving in line at the voting station, and plan to call a colleague for help if it gets too busy. Let X be a random variable describing the number of voters that arrive at the station over a 15 minute period.

\sim

$X \sim \text{Pois}(\lambda)$ \lambda

# Geometric distribution

- **Canonical Example:** Suppose you flip the same biased coin repeatedly. How many times do you flip the coin before you see your first Heads?

Coin has $Pr(H) = p$

Each flip is $Ber(p)$

$P(X=1) = p$

$P(X=2) = P(\text{heads on } 2^{nd} \mid \text{tails on } 1^{st} \text{ flip})\, P(\text{tails on } 1^{st} \text{ flip}) = p(1-p)$

$$\boxed{P(X=k) = p(1-p)^{k-1}}$$

# Geometric distribution

- **Definition**: A discrete random variable $x$ has a geometric distribution with parameter p, where $0 \leq p \leq 1$ , if its probability mass function is given by

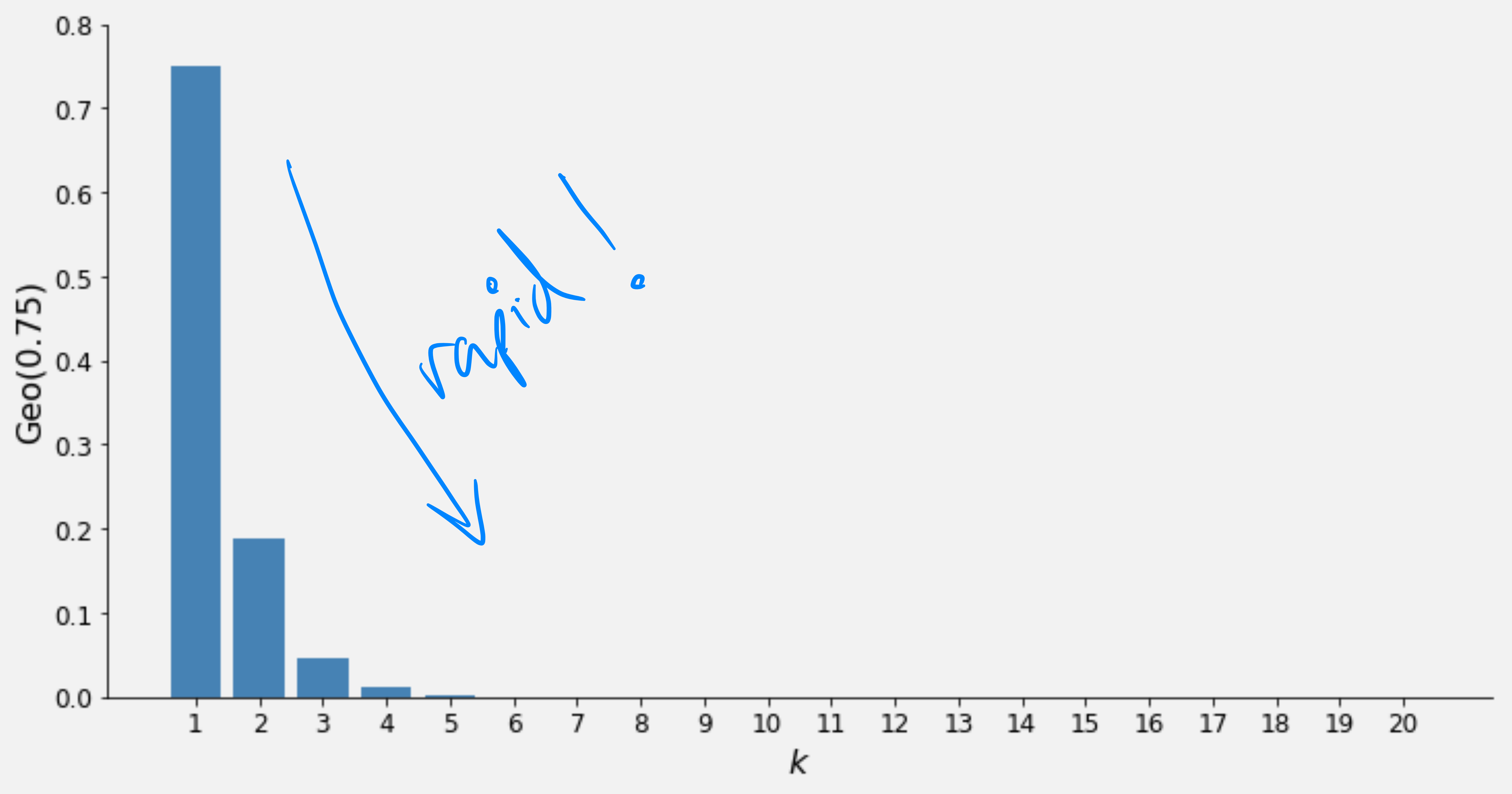
$$p_X(k) = P(X = k) = (1-p)^{k-1}p \quad \text{for} \quad k = 1, 2, 3, \ldots$$

$X \sim Geo(p)$

# Geometric distribution

- **Definition**: A discrete random variable $X$ has a geometric distribution with parameter p, where $0 \leq p \leq 1$ , if its probability mass function is given by

$$p_X(k) = P(X = k) = (1-p)^{k-1}p \quad \text{for} \quad k = 1, 2, 3, \ldots$$

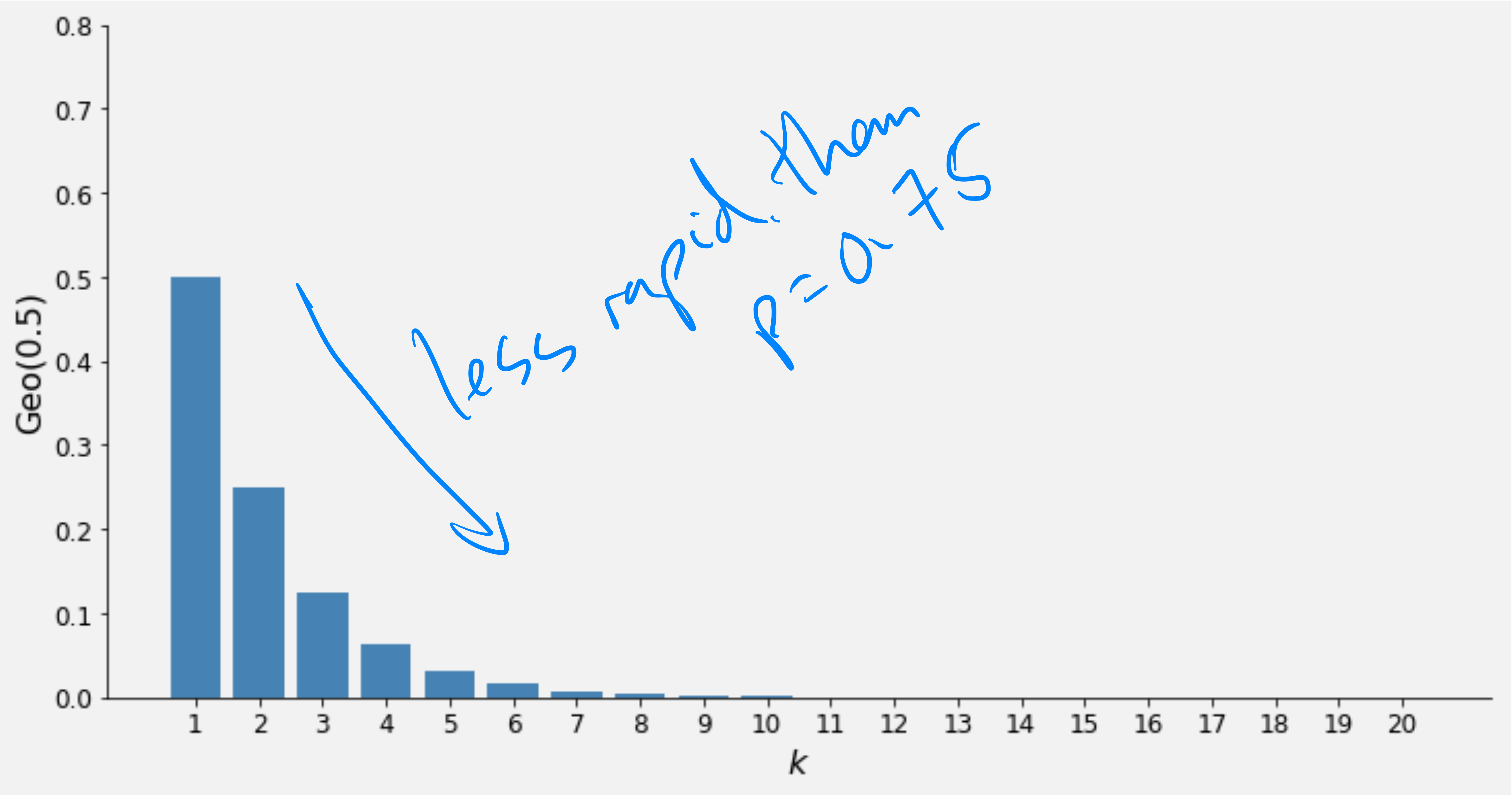



$X \sim Geo(p)$

# Geometric distribution

- **Definition**: A discrete random variable $x$ has a geometric distribution with parameter p, where $0 \leq p \leq 1$ , if its probability mass function is given by

$$p_X(k) = P(X = k) = (1-p)^{k-1}p \quad \text{for} \quad k = 1, 2, 3, \ldots$$




$X \sim Geo(p)$

# Geometric distribution

- **Assumptions**: What did we assume when writing down the geometric distribution?
- Each flip is Ber(p)
- Flips are independent of each other.

# Negative binomial distribution

- **Canonical Example:** Suppose you flip the same biased coin repeatedly. How many times do you flip the coin before you see three (or *r* ) Heads?

If I get $r$ Heads

Waited $k$ flips total

$\Rightarrow$ $k-r$ Tails

r Heads

k-r Tails

$$(\#)\,p^r\,(1-p)^{k-r}$$

???

$$\underbrace{p\,(1-p)(1-p)\,p\,p\,p\,p\,(1-p)\,p\,(1-p)\,p}_{\text{lots of orders!}}$$

# Negative binomial distribution

- **Canonical Example:** Suppose you flip the same biased coin repeatedly. How many times do you flip the coin before you see ~~three~~ $r$ Heads?

Total of $k$ flips

$r$ Heads

$k-r$ Tails

last flip heads!

$r-1$ in the bins

$\# \binom{k-1}{r-1}$

$$\underbrace{\_\_\ \_\_\ \_\_\ \cdots\ \_\_\ \_\_}_{k-1 \text{ "bins"}}\ \underset{k}{\underline{H}}$$

$1 \quad 2 \quad 3 \quad \ldots \quad k-2 \quad k-1 \quad k$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, *r>1* and p, $0 \le p \le 1$ if its probability mass function is given by $$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$

r=1 → NB becomes Geometric

r = # H

k = # H + # T

# T = k - # H

Wik... says diff! ??

Be careful about parametrization.

$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, $r>1$ and p, $0 \leq p \leq 1$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$

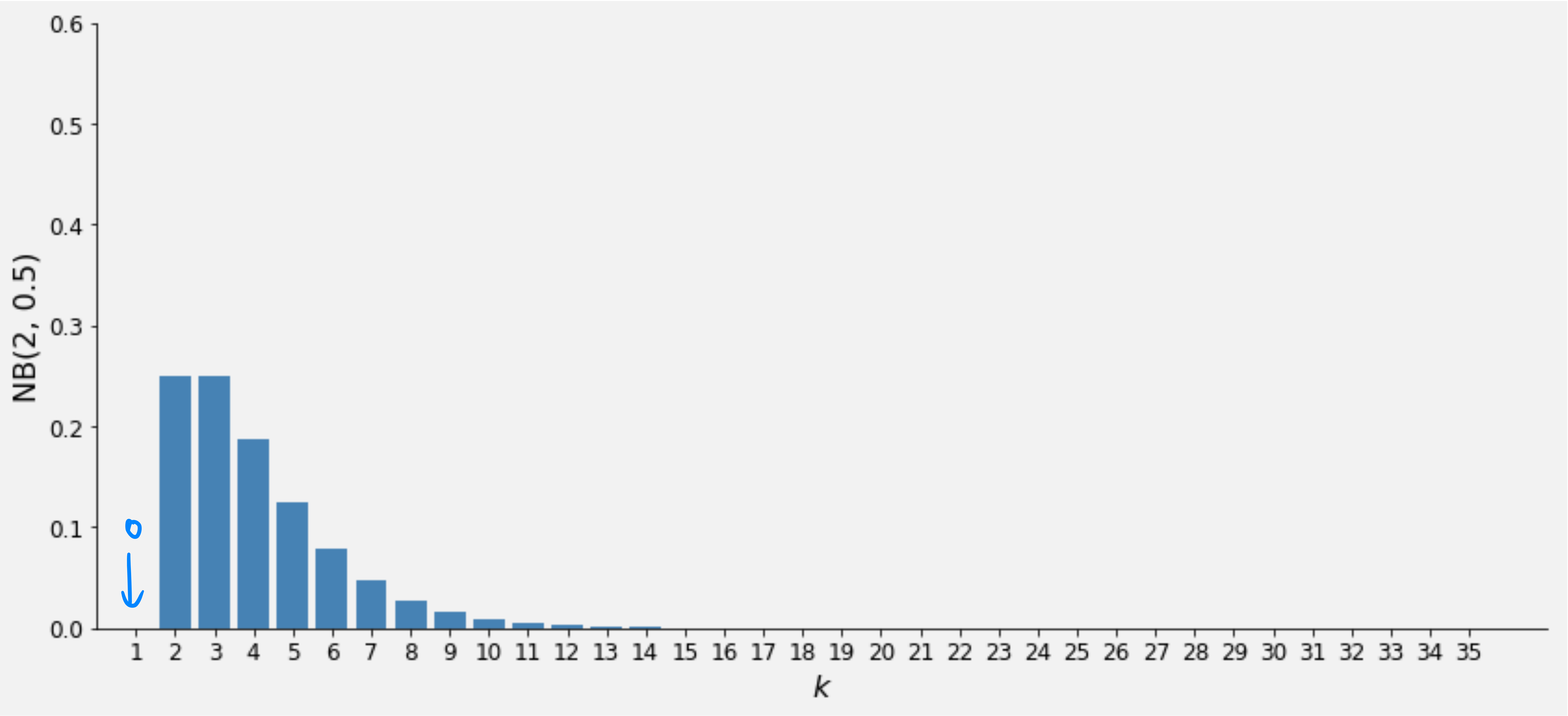

$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, *r>1* and p, $0 \leq p \leq 1$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$

$r = 2$

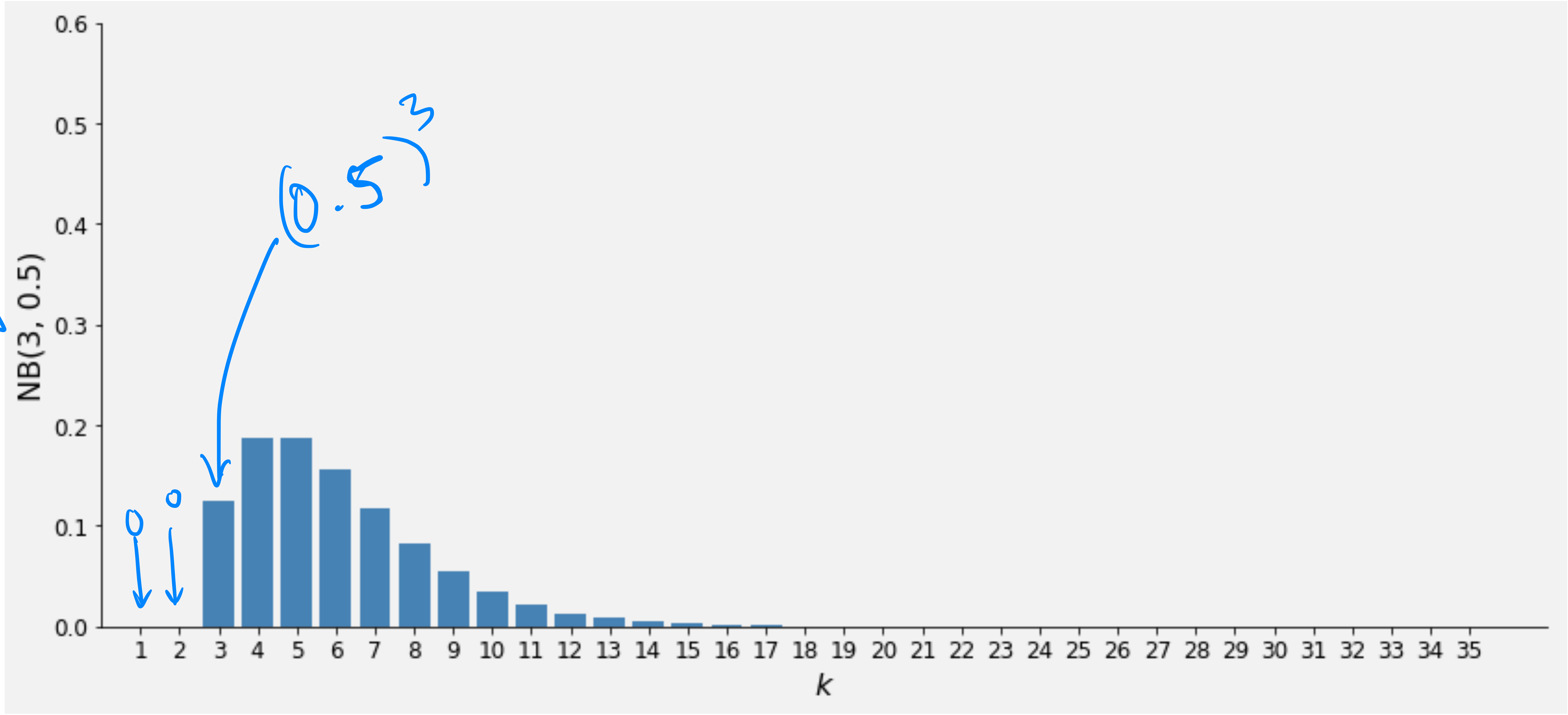



$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, *r>1* and p, $0 \leq p \leq 1$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$

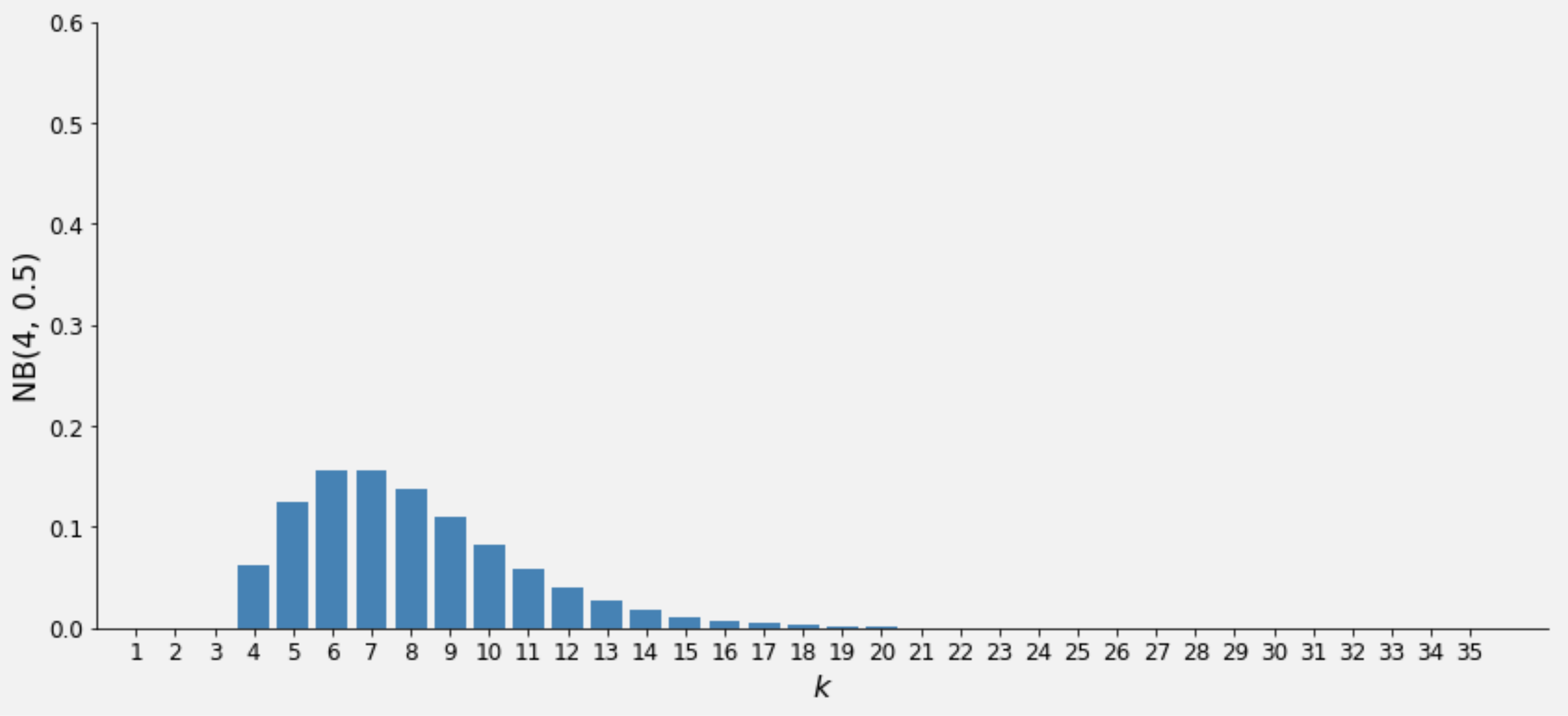



$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, $r>1$ and p, $0 \leq p \leq 1$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$




$X \sim NB(r, p)$

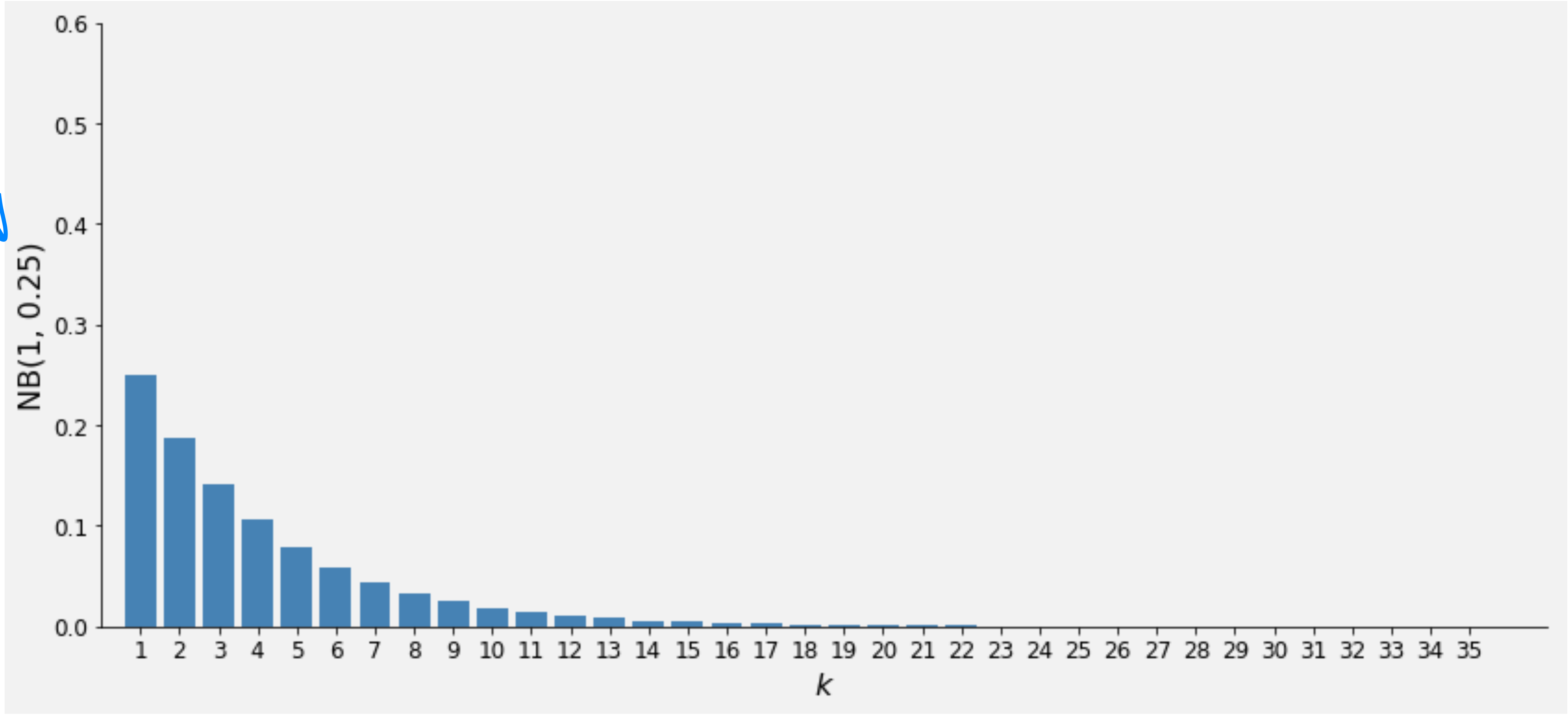
# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, $r>1$ and p, $0 \leq p \leq 1$ if its probability mass function is given by

$$p_X(k) = P(X=k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$

$p=0.25$

$r=1$

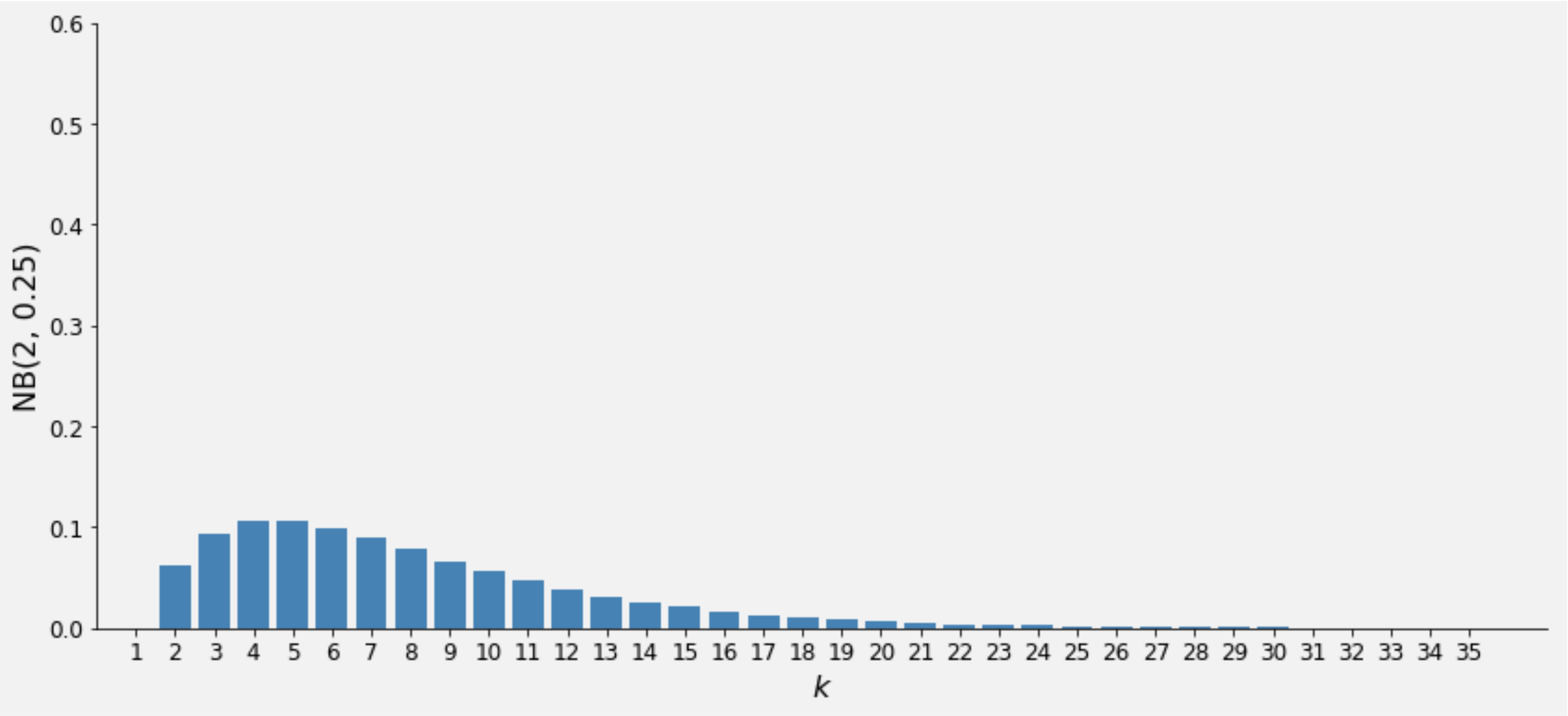



$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, $r>1$ and p, $0 \leq p \leq 1$ if its probability mass function is given by

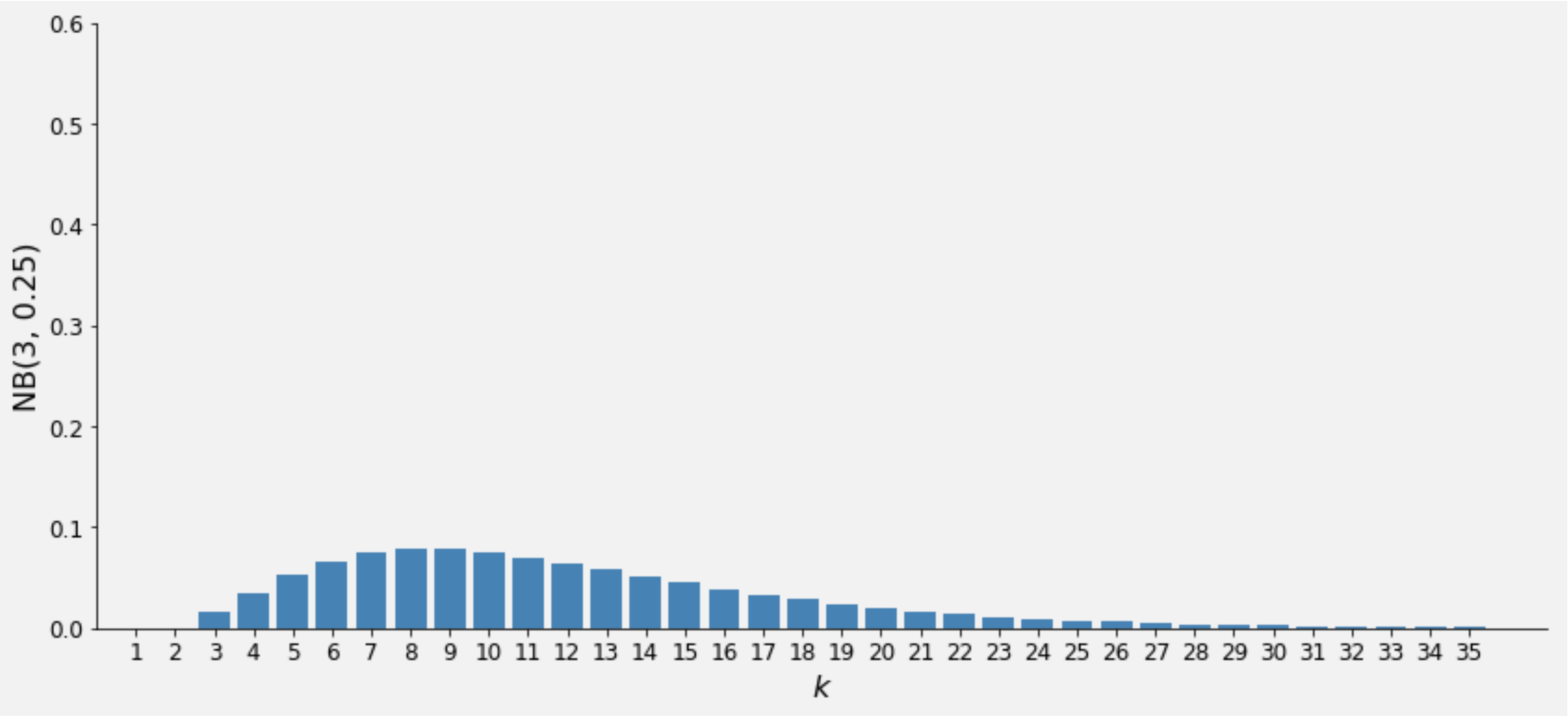
$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$




$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, *r>1* and p, $0 \leq p \leq 1$ if its probability mass function is given by $$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$

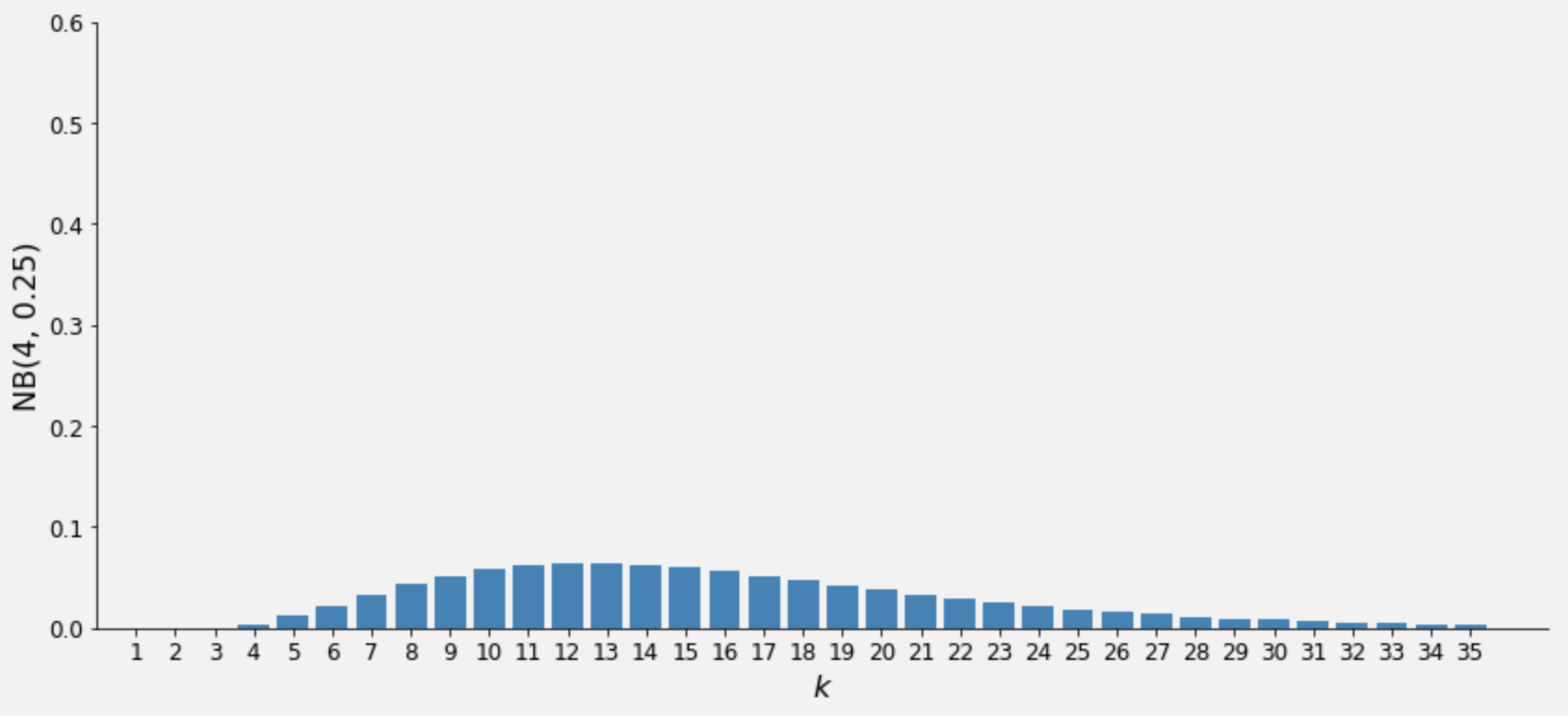



$X \sim NB(r, p)$

# Negative binomial distribution

- **Definition**: A discrete random variable $x$ has a negative binomial distribution with parameters r, $r>1$ and p, $0 \leq p \leq 1$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad \text{for} \quad k = r, r+1, \ldots$$




$X \sim NB(r, p)$

# Negative Binomial

- **Assumptions**: What did we assume when writing down the negative binomial distribution?
- Each flip Ber(p)
- Flips are indep.

# Poisson distribution

- **Canonical Example:** You get texts during class :/ !! at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

LIMITS!

Binomial: discrete trials that can "succeed" "fail"

1 H

0 T

rate $\frac{\text{texts}}{\text{hr}} = \lambda = n \cdot p$

$n$: $\frac{\text{flips}}{\text{hr}}$, $p$: $\frac{\text{text}}{\text{flip}}$

$\lambda = np$

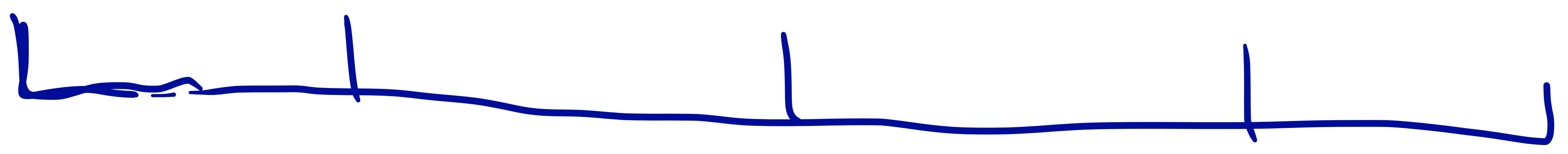
# Poisson distribution

- **Canonical Example:** You get texts during class :/ !! at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

rate $\lambda = np$

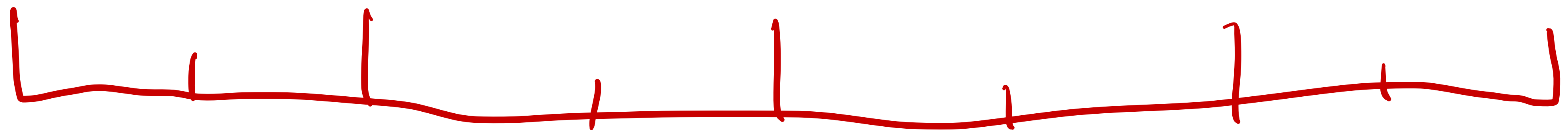
$\lambda = 4(0.5)$

$\lambda = 2$



1 hr

fix $\lambda$ send $n \to \infty$



1 hr

$\lambda = 2 = 8(0.25)$

# Poisson distribution

- **Canonical Example:** You get texts during class :/ !! at a rate of 29 texts per hour. What is the probability that you get 29 texts in an hour? 12 texts in an hour? 107 texts in an hour?

fix $\lambda = np$, send $n \to \infty$ $\quad n = \frac{\lambda}{p}$ or $p = \frac{\lambda}{n}$

$$B(n,p) \sim P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$\frac{\lambda^k}{k!} \lim_{n\to\infty} \cancel{\frac{n!}{(n-k)!} \left(\frac{1}{n}\right)^k} \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$\downarrow e^{-\lambda}$$

$$\frac{n!}{(n-k)!} \cdot \frac{1}{n^k} = \frac{n(n-1)(n-2)\cdots}{(n-k)(n-k-1)(n-k-2)\cdots} \cdot \frac{1}{n^k}$$

$$= \frac{n \quad (n-1) \quad (n-2) \cdots (n-k+1)}{n \quad n \quad n \cdots n}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-k+1}{n} \longrightarrow 1$$

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e$$

# Poisson distribution

- **Definition**: A discrete random variable $x$ has a Poisson distribution with parameter $\lambda$ , $\lambda > 0$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

rate = $\lambda$

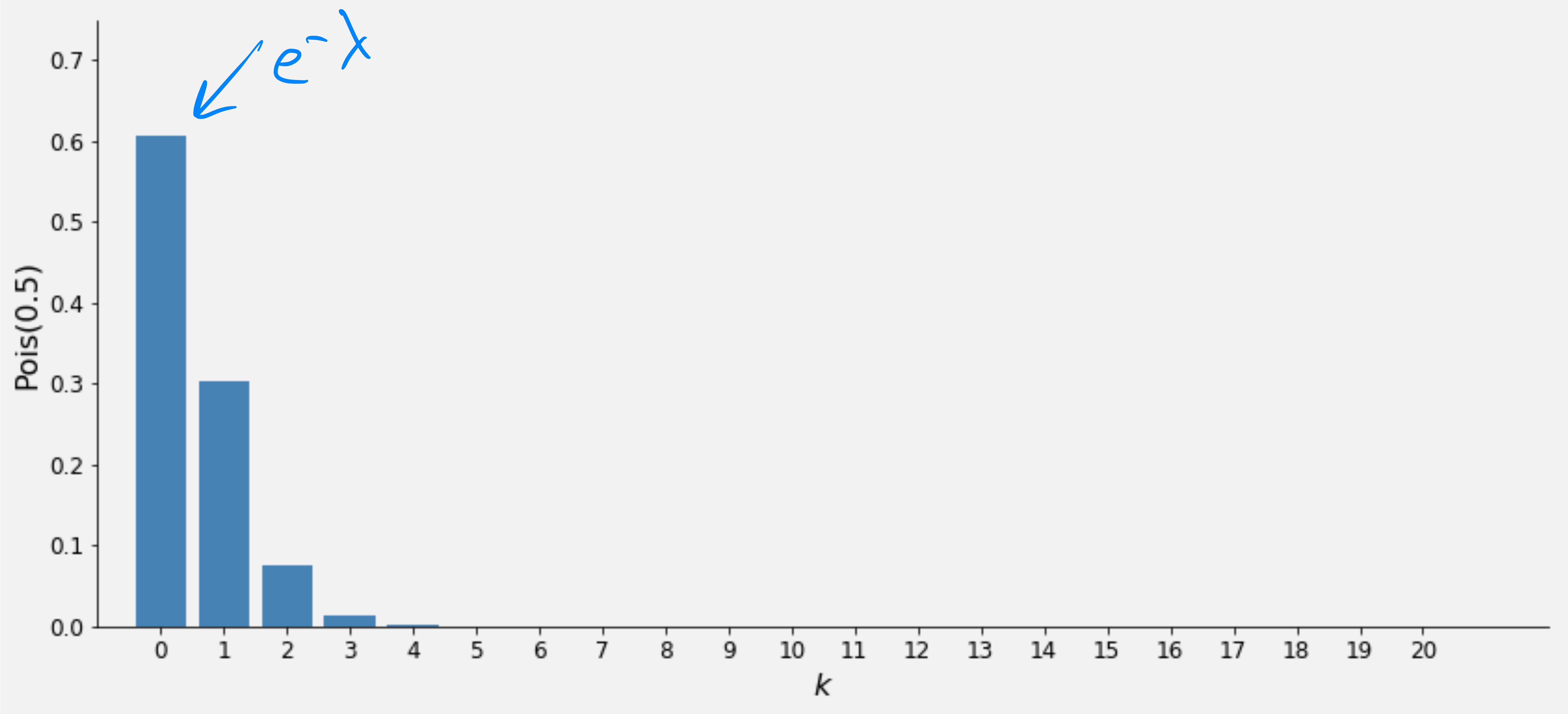
# testing (random) $k$

$X \sim Pois(\lambda)$

# Poisson distribution

- **Definition**: A discrete random variable $x$ has a Poisson distribution with parameter $\lambda$ , $\lambda > 0$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

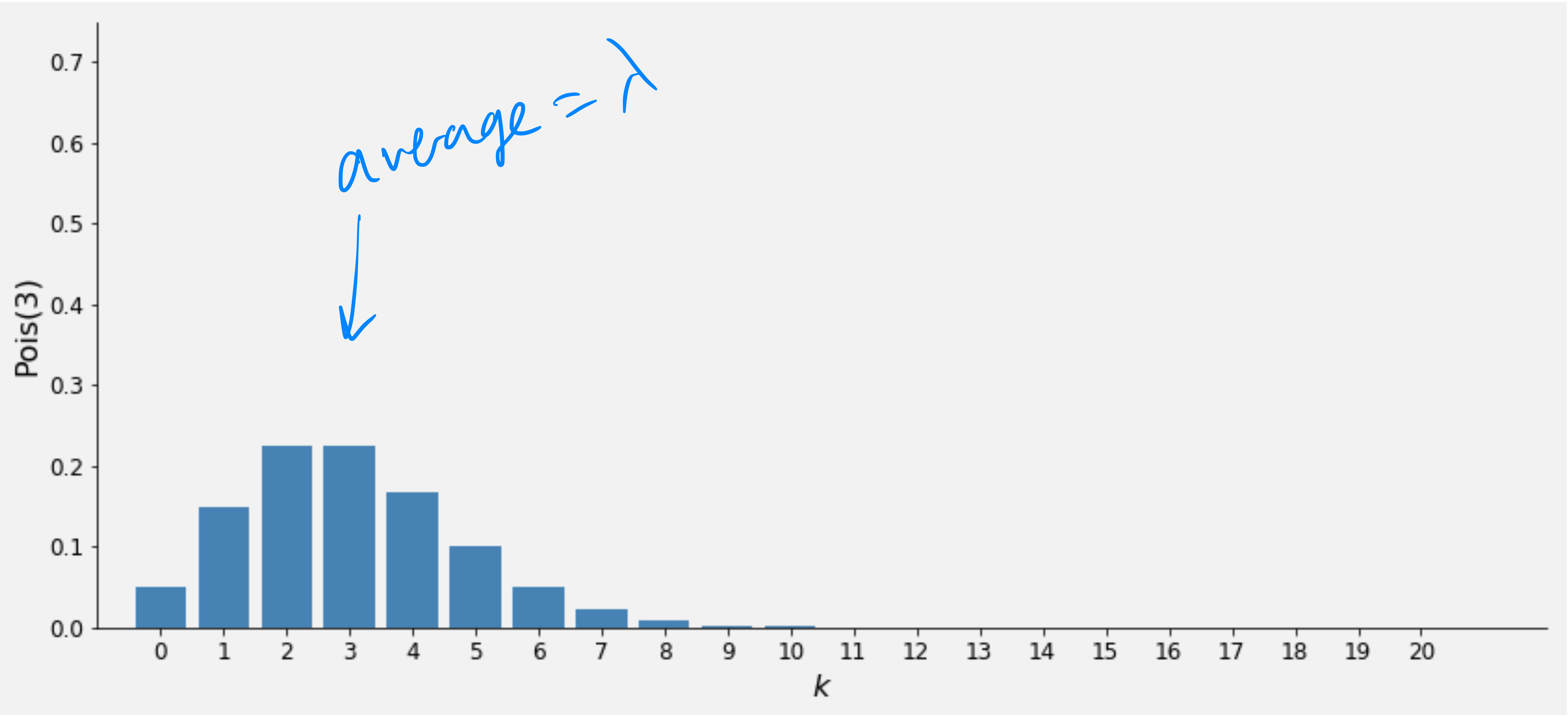



$X \sim Pois(\lambda)$

# Poisson distribution

- **Definition**: A discrete random variable $x$ has a Poisson distribution with parameter $\lambda$ , $\lambda > 0$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

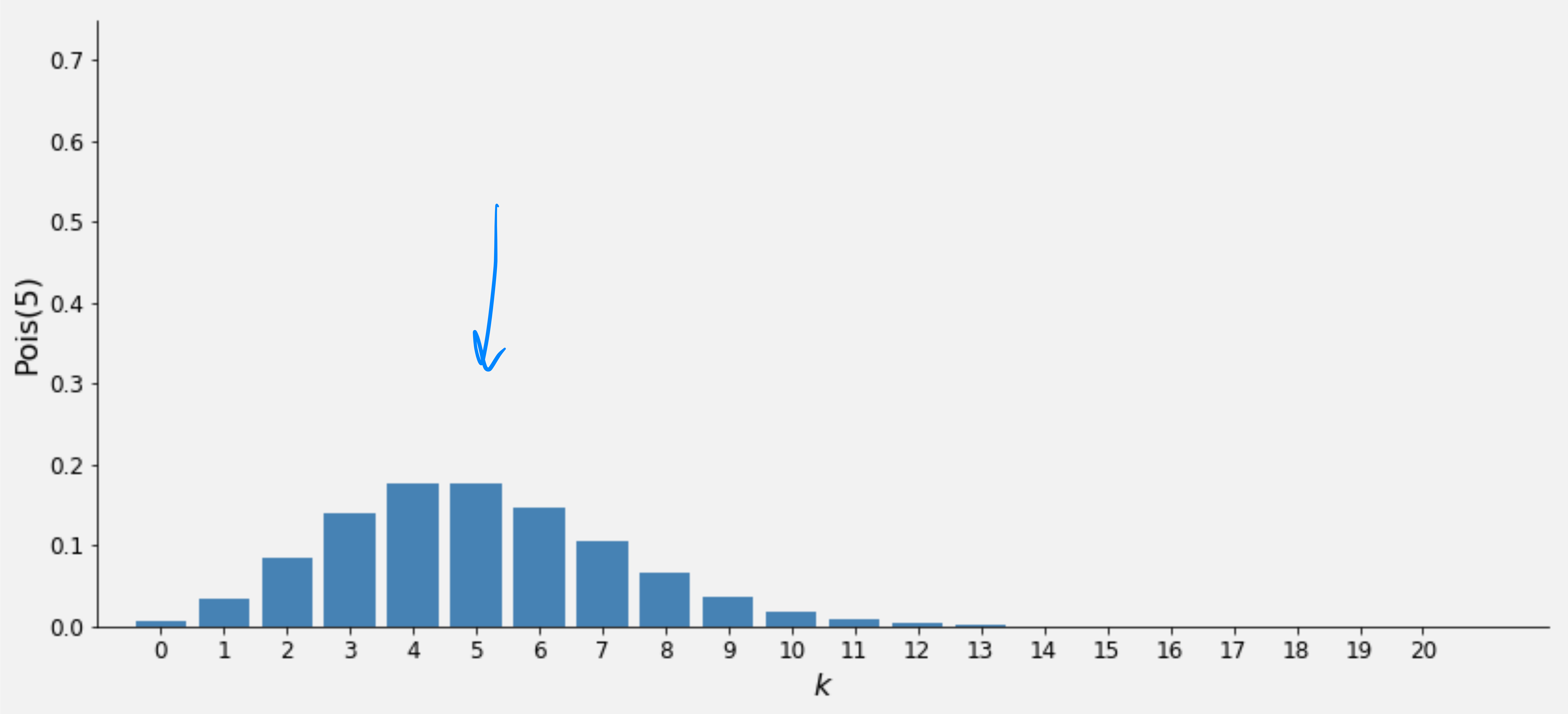



$X \sim Pois(\lambda)$

# Poisson distribution

- **Definition**: A discrete random variable $x$ has a Poisson distribution with parameter $\lambda$ , $\lambda > 0$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

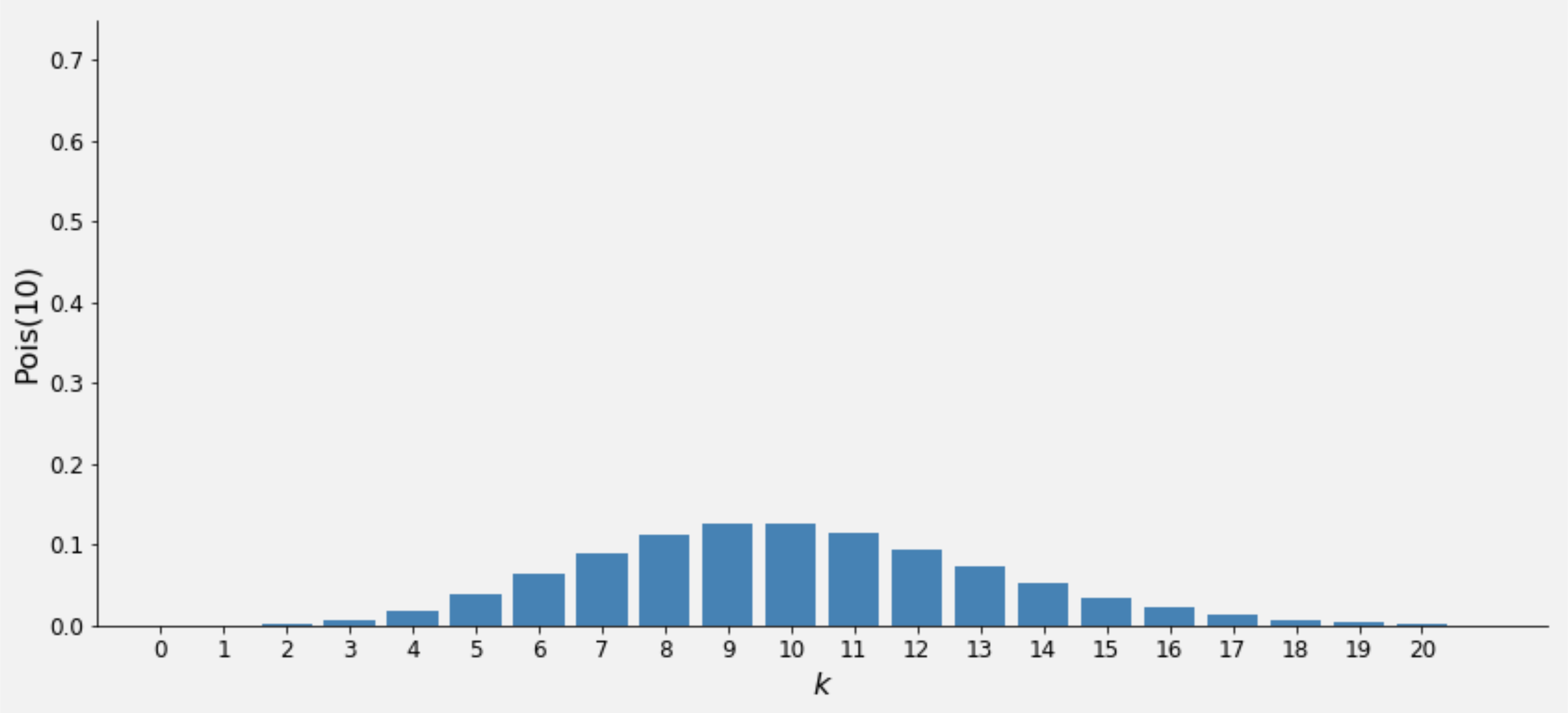



$X \sim Pois(\lambda)$

# Poisson distribution

- **Definition**: A discrete random variable $x$ has a Poisson distribution with parameter $\lambda$ , $\lambda > 0$ if its probability mass function is given by

$$p_X(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for} \quad k = 0, 1, 2, \ldots$$




$X \sim Pois(\lambda)$

# Poisson

- **Assumptions**: What did we assume when writing down the Poisson distribution?
- Each "time" interval was indep. of others.
- Rate $\lambda$ is const

# Poisson with a twist?

- **Question**: Question: Suppose arrivals are described well by a Poisson random variable. What is the probability that the time between two arrivals is between 5 and 10 minutes?

? ? ?

Exponential.