COMP 182 Algorithmic Thinking

Discrete Probability

Luay Nakhleh
Computer Science
Rice University

Reading Material

- Chapter 7, Sections 1-4

Experiment, Sample Space, and Event

- An experiment is a procedure that yields one of a given set of possible outcomes (tossing a coin, rolling a die five times, etc.).
- The sample space of the experiment is the set of possible outcomes (for the coin toss experiment, the sample space is {H,T}).
- An event is a subset of the sample space (the possible events that correspond to the coin toss experiment are {}, {H}, {T}, {H,T}).

Experiment, Sample Space, and Event

- For each of the following, what is the sample space? What is an example of an event?
  - Experiment 1: Rolling a 6-sided die twice.
  - Experiment 2: Generating an Erdos-Renyi graph with parameters n and p.

Assigning Probabilities

- Let S be a sample space of an experiment with finite number of outcomes. We assign a probability $p(s)$ to every outcome s, so that

  1. $0 \leq p(s) \leq 1$ for each $s \in S$, and
  2. $\sum_{s \in S} p(s) = 1.$

Assigning Probabilities

- The function p from the set of all outcomes of the sample space S is called a probability distribution.

Uniform Distribution

- The <u>uniform distribution</u> on a set S with n elements assigns probability $1/n$ to each element of S (all n elements are equally likely).

Probability of an Event: Laplace's Definition

Pierre-Simon Laplace

- **If** S is a finite sample space of **equally likely outcomes**, and E is an event, that is, a subset of S, then the probability of E is $p(E) = |E| / |S|$.

Probability of an Event

- Back to the experiment of rolling a die twice:
  - What is the probability of the event “the sum of the two numbers is 8”?

Probability of an Event

- When the elements of the sample space are not all equally likely, Laplace's definition doesn't work!
- The probability of an event E is then defined in terms of the probabilities of its elements.

Probability of an Event

- The probability of the event E is

$$p(E) = \sum_{s \in E} p(s)$$

Probability of an Event

- Back to the experiment of ER graphs with n and p:
  - What is the probability of the event “the generated graph has exactly k edges”?

Combinations of Events

- If $E_1, E_2, \ldots$ is a sequence of pairwise disjoint events in a sample space S, then

$$p\left(\bigcup_i E_i\right) = \sum_i p(E_i)$$

Conditional Probability

- Let E and F be two events with $p(F)>0$. The conditional probability of E given F, denoted by $p(E \mid F)$, is defined as

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Independence

- The events E and F are <u>independent</u> if and only if $p(E \cap F) = p(E)p(F)$.
- For example, consider the space of randomly generated bit strings of length four (all 16 have the same probability), and consider:
    - *E*: the string begins with 1
    - *F*: the string contains an even number of 1s.

Pairwise and Mutual Independence

- The events $E_1, E_2, \ldots, E_n$ are <u>pairwise independent</u> if and only if $p(E_i \cap E_j) = p(E_i)p(E_j)$ for all pairs i and j with $i \neq j$.
- The events $E_1, E_2, \ldots, E_n$ are <u>mutually independent</u> if and only if for every subset $E' \subseteq \{E_1, E_2, \ldots, E_n\}$ with $|E'| \geq 2$ we have

$$p\left(\bigcap_{E_i \in E'} E_i\right) = \prod_{E_i \in E'} p(E_i)$$

- Mutually independent events are also pairwise independent events.
- Pairwise independent events are not necessarily mutually independent events.
  - Experiment: Toss a fair coin twice.
  - Events:
    - E1: The first toss is H.
    - E2: The second toss is H.
    - E3: Both tosses give the same outcome.

Bernoulli Trials

- Suppose an experiment can have only two possible outcomes, e.g., the flipping of a coin or the random generation of a bit (recall the random ER graphs?).
- Each performance of the experiment is called a Bernoulli trial.
- One outcome is called a success and the other a failure.
- If p and q are the probabilities of success and failure, respectively, then $p+q=1$.

The Binomial Distribution B(k:n,p)

- The probability of exactly *k* successes in *n* independent Bernoulli trials, with probability of success *p* and probability of failure *q=1-p*, is

$$\binom{n}{k} p^k q^{n-k}$$

Bayes' Theorem

Bayes' Theorem

- Suppose that E and F are events from a sample space S such that $p(E) \neq 0$ and $p(F) \neq 0$. Then:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Bayes' Theorem

- Suppose that E and F are events from a sample space S such that $p(E) \neq 0$ and $p(F) \neq 0$. Then:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

$p(E)$

- Suppose that one person in 100,000 has a particular disease. There is a test for the disease that gives a positive result 99% of the time when given to someone with the disease. When given to someone without the disease, 99.5% of the time it gives a negative result. Find

    1. the probability that a person who tests positive has the disease.
    2. the probability that a person who tests negative does not have the disease.

Generalized Bayes' Theorem

- Suppose that $\{F_1, F_2, \ldots, F_n\}$ is a partition of the sample space S. Then, for an event E, we have

$$p(F_j \mid E) = \frac{p(E \mid F_j)p(F_j)}{\sum_{i=1}^{n} p(E \mid F_i)p(F_i)}$$

Random Variables, Expected Value and Variance

Random Variables

- A random variable is a function from the sample space of an experiment to the set of real numbers.
- A coin is tossed twice. Let *X(t)* be the random variable that equals that number of heads that appear when t is the outcome. Then *X(t)* takes on the following values:
  - $X(HH)=2$
  - $X(HT)=X(TH)=1$
  - $X(TT)=0$

Random Variables

- The distribution of a random variable X on a sample space S is the set of pairs $(r,p(X=r))$ for all $r \in X(S)$, where $p(X=r)$ is the probability that X takes the value r.
- Example: A fair coin is tossed twice and $X(t)$ is the number of heads in outcome t. The distribution of X is
    - {(0,0.25),(1,0.5),(2,0.25)}

Random Variables:
Illustration on Random Graphs

- Experiment: Given a set of n nodes, for every set of two nodes, connect them with an edge with probability p (recall Erdos-Renyi?)
- What is the sample space? What is its size?
- Define a random variable $X(g)$ that equals the number of edges in g.
- What are the possible values of $X(g)$?
- What is the probability distribution of $X(g)$?

Random Variables: Illustration on Random Graphs

- Experiment: Given a set of *n* nodes, for every set of two nodes, connect them with an edge with probability *p* (recall Erdos-Renyi?)
- What is the sample space? What is its size?
  - *S* is the set of all graphs with *n* nodes.
  - *S* contains the set of all *n*-node graphs with *0* edges, *1* edge, 2 edges,…, *n(n-1)*/2 edges. So, the size of *S* is

$$|S| = \sum_{k=0}^{n(n-1)/2} \binom{n(n-1)/2}{k} = 2^{n(n-1)/2}$$

Random Variables: Illustration on Random Graphs

- Experiment: Given a set of n nodes, for every set of two nodes, connect them with an edge with probability p (recall Erdos-Renyi?)
- Define a random variable $X(g)$ that equals the number of edges in g.
- What are the possible values of $X(g)$?
  - $X(g) \in \{0,1,\ldots,n(n-1)/2\}$.

Random Variables: Illustration on Random Graphs

- Experiment: Given a set of n nodes, for every set of two nodes, connect them with an edge with probability p (recall Erdos-Renyi?)
- Define a random variable $X(g)$ that equals the number of edges in g.
- What is the probability distribution of $X(g)$?

$$p(X(g)=k) = \binom{n(n-1)/2}{k} p^k (1-p)^{\frac{n(n-1)}{2}-k} \quad k \in \left\{0, 1, \ldots, \frac{n(n-1)}{2}\right\}$$

Expected Values

- The expected value (also called the expectation or mean) of a (discrete) random variable X on the sample space S is

$$\mathbb{E}(X) = \sum_{s \in S} p(s) X(s)$$

Expected Values

- Example: A fair coin is tossed twice and *X(t)* is the number of heads.

$$\begin{aligned}\mathbb{E}(X) &= p(HH)\cdot X(HH)+p(HT)\cdot X(HT)+p(TT)\cdot X(TT)+p(TH)\cdot X(TH)\\ &= 0.25\cdot 2+0.25\cdot 1+0.25\cdot 0+0.25\cdot 1\\ &= 1\end{aligned}$$

Expected Values

- What is the expected number of edges in a random graph with *n* nodes and probability *p* (using the ER procedure)?
  - Based on the results we saw before, we have

$$\begin{aligned}\mathbb{E}(X(g)) &= \sum_{k=0}^{n(n-1)/2} k \cdot p(X(g) = k) \\ &= \sum_{k=0}^{n(n-1)/2} k \binom{n(n-1)/2}{k} p^k (1-p)^{\frac{n(n-1)}{2} - k} \\ &= \frac{n(n-1)}{2} p\end{aligned}$$

Linearity of Expectations

- If X_i, i=1,2,…,n, are random variables on S, and if a and b are real numbers, then

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n)$$

$$\mathbb{E}(aX_i + b) = a\mathbb{E}(X_i) + b$$

Average-case Analysis

- What is the average-case running time of **LinearSearch** (for finding whether an element x exists in an array of n elements)?
- Assume x is in the input array with probability p.
- Assume that if x is in the array, it can be in any position with equal probability.
- What random variable do we define? What is the expected value of the variable?

The Geometric Distribution

- So far, we have discussed random variables with a finite number of possible outcomes.
- Consider the following experiment: A coin with probability of tails being p is tossed repeatedly until it comes up tails. What is the expected number of tosses until this coin comes up tails?
- The sample space here is {T,HT,HHT,HHHT,…}, which is infinite.

The Geometric Distribution

- Let X be the random variable equal to the number of tosses in an element in the sample space.
- We have $p(X=k)=(1-p)^{k-1}p$.
- Then,

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} kp(X = k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = p\sum_{k=1}^{\infty} k(1-p)^{k-1} = p\frac{1}{p^2} = \frac{1}{p}$$

The Geometric Distribution

- Let X be the random variable equal to the number of tosses in an element in the sample space.
- We have $p(X=k)=(1-p)^{k-1}p$.
- Then,

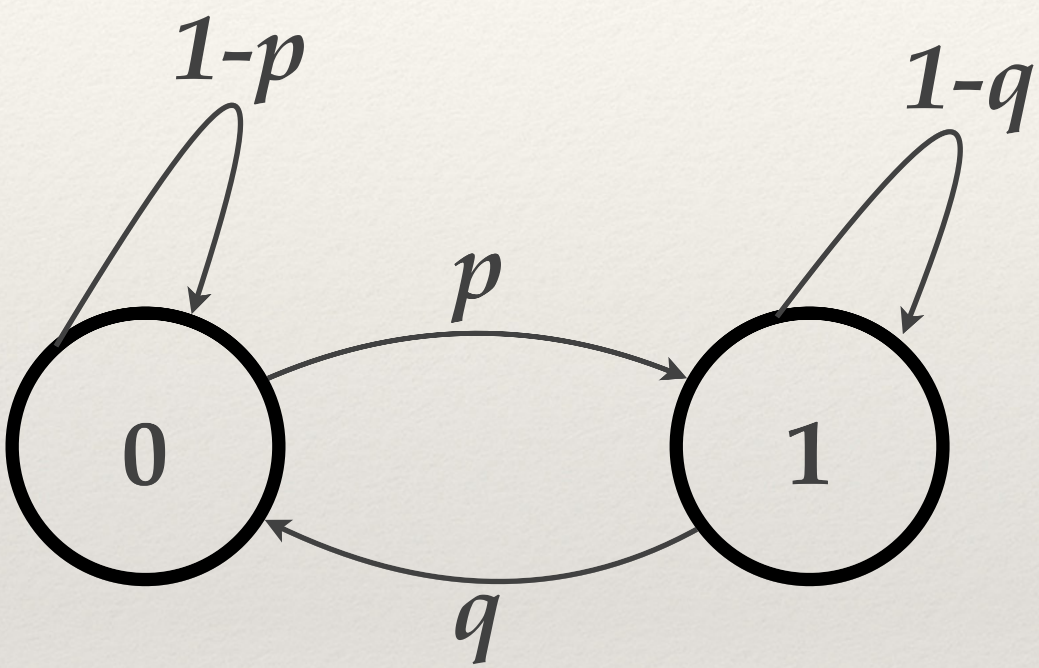
$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2} \quad (|x| < 1)$$

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} kp(X=k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = p\sum_{k=1}^{\infty} k(1-p)^{k-1} = p\frac{1}{p^2} = \frac{1}{p}$$

The Geometric Distribution

- A random variable X has a geometric distribution with parameter p if $p(X=k)=(1-p)^{k-1}p$ for $k=1,2,3,\ldots$, where p is a real number with $0 \leq p \leq 1$.
- If $X \sim Geometric(p)$, then $\mathbb{E}(X) = 1/p$.

The Geometric Distribution



How is the duration in state 0 distributed?

Independent Random Variables

- The random variables X and Y on a sample space S are independent if

$$p(X = x \text{ and } Y = y) = p(X = x)p(Y = y)$$

Independent Random Variables

- If random variables X and Y on sample space S are independent, then

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Variance

- The expected value of a random variable tells us its average value, but nothing about how widely its values are distributed.
- Contrast X and Y on $S=\{1,2,3,4,5,6\}$, where
  - $X(s)=0$ for all $s \in S$
  - $Y(s)=-1$ for $s \in \{1,2,3\}$ and $Y(s)=1$ for $s \in \{4,5,6\}$

Variance

- Let *X* be a random variable on a sample space *S*. The variance of *X*, denoted by *V(X)*, is

$$V(X) = \sum_{s \in S} (X(s) - \mathbb{E}(X))^2 p(s)$$

- The standard deviation of X, denoted by σ(X), is defined to be

$$\sigma(X) = \sqrt{V(X)}$$

Variance

$$
\begin{aligned}
V(X) &= \textstyle\sum_{s\in S}(X(s)-\mathbb{E}(X))^2 p(s) \\
&= \textstyle\sum_{s\in S} X(s)^2 p(s) - 2\mathbb{E}(X)\sum_{s\in S} X(s)p(s) + \mathbb{E}(X)^2 \sum_{s\in S} p(s) \\
&= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\
&= \mathbb{E}(X^2) - \mathbb{E}(X)^2
\end{aligned}
$$

Variance

- If X is a random variable on a sample space S and $\mathbb{E}(X) = \mu$, then

$$V(X) = \mathbb{E}((X - \mu)^2)$$

Bienayme's Formula

- If X_i, $i=1,2,\ldots,n$, are pairwise independent random variables on S, then

$$V(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X_i)$$

Variance

- If X is a random variable and c is a constant then,

$$V(cX) = c^2 V(X)$$

Bienayme's Formula

- If X_i, i=1,2,…,n, are random variables (not necessarily independent) on S, then

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i=1}^{n}\sum_{j>i} Cov(X_i, X_j)$$

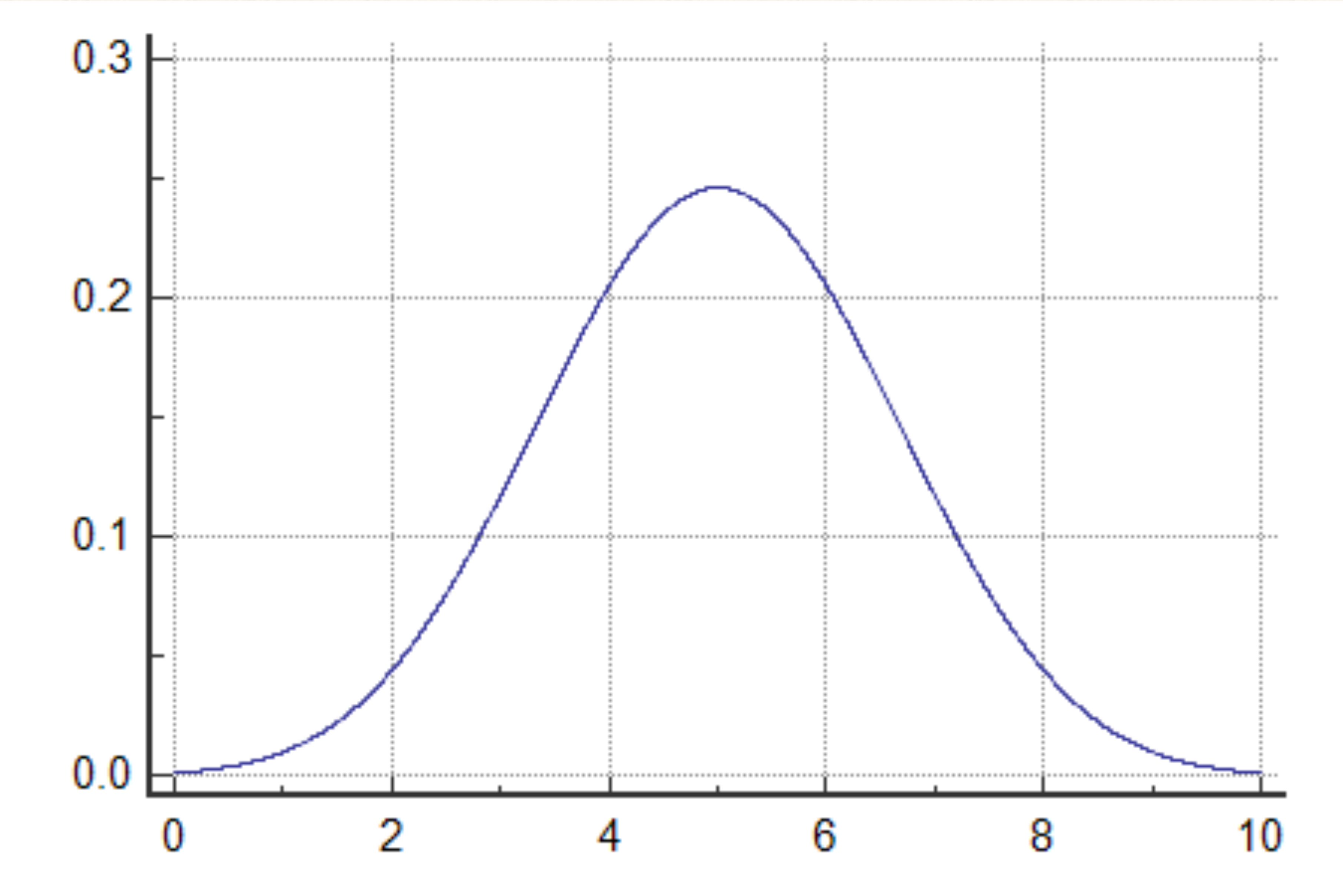
where

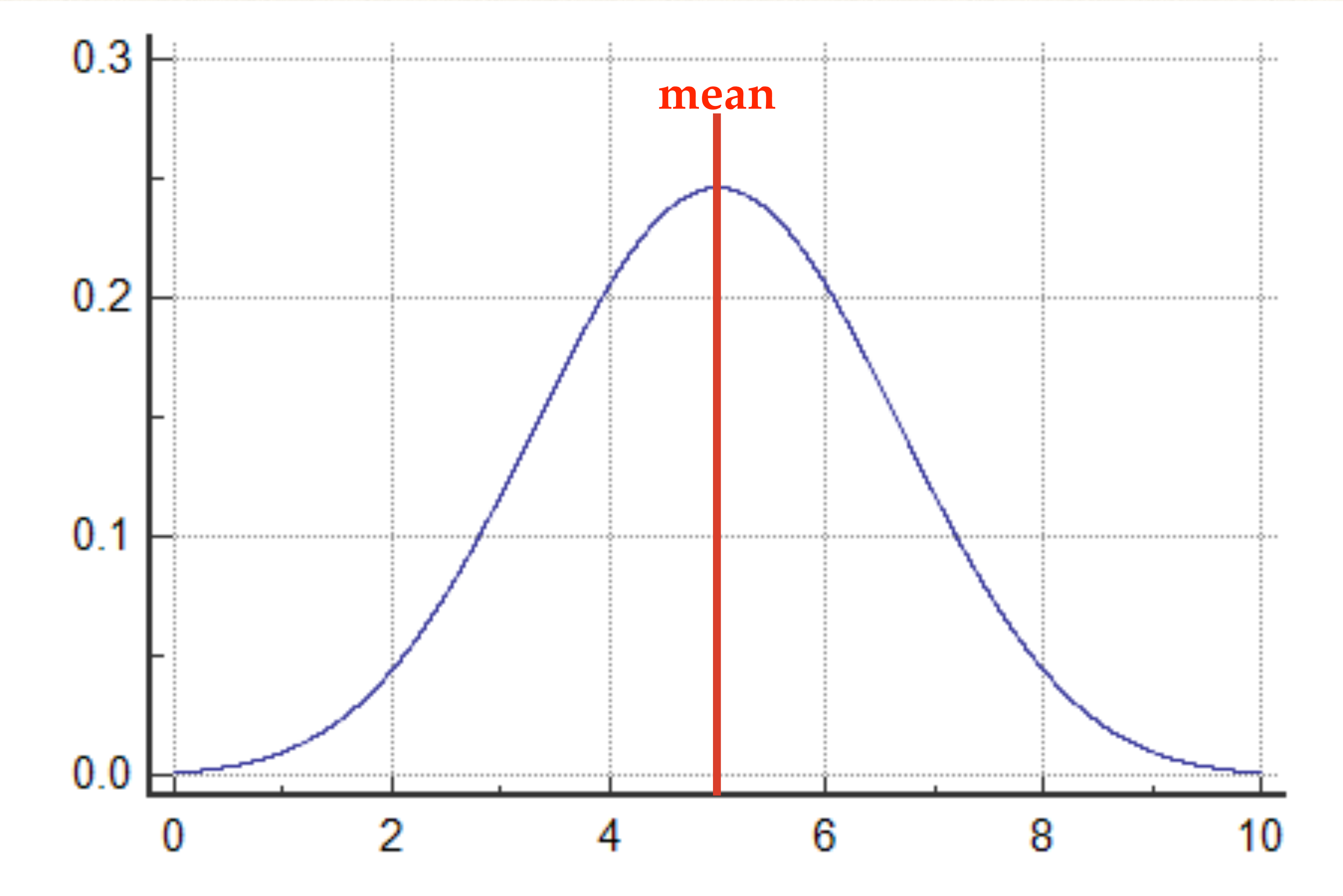
$$Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Tail Bounds

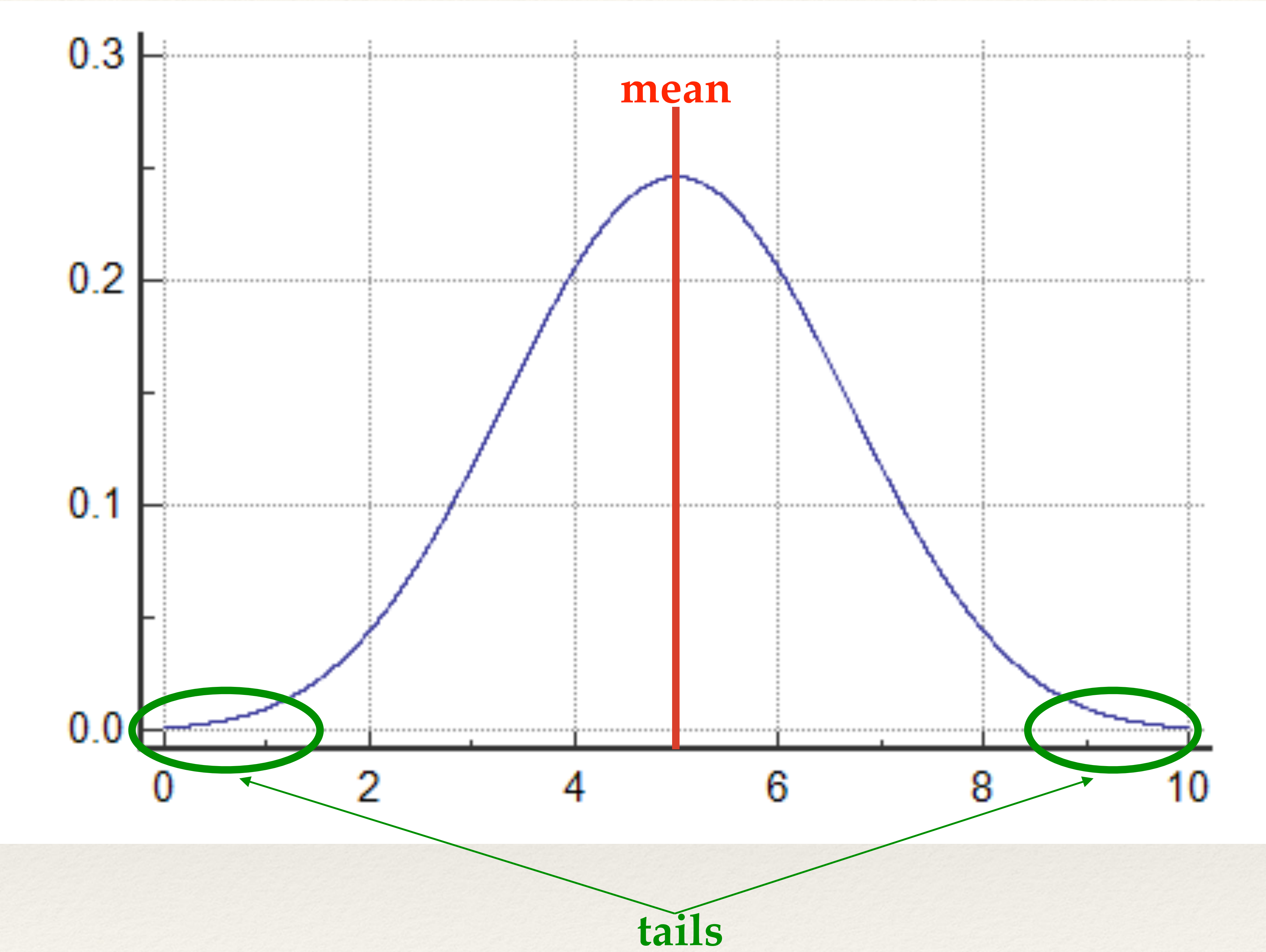
What Is This About?

- How large can a random variable get?
- In other words, how far from the mean can a value that the random variable takes be?

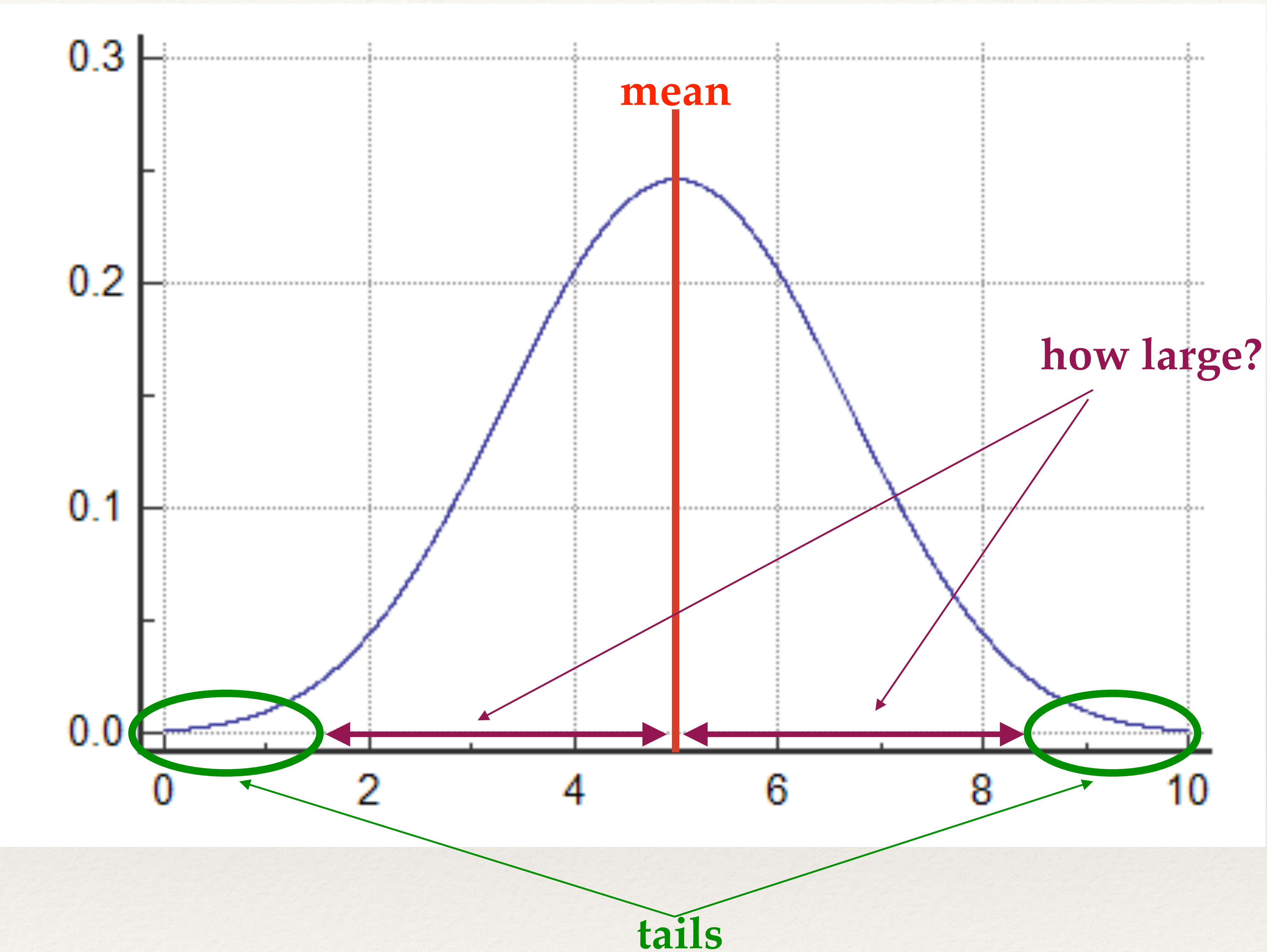

0.3
0.2
0.1
0.0
0
2
4
6
8
10



mean
0.3
0.2
0.1
0.0
0
2
4
6
8
10



0.3
0.2
0.1
0.0
mean
0
2
4
6
8
10
tails


0.3
0.2
0.1
0.0
0
2
4
6
8
10
mean
how large?
tails

Why Do We Care?

- Example:
  - X is the number of steps an algorithm takes.
  - $\mathbb{E}(X)$ is the average-case running-time of the algorithm.
  - Can the algorithm, on average, take $2n$ steps, but on some inputs take, say, $500n^2$ steps?

Markov's Inequality

- Let X be a random variable that takes only nonnegative values. Then, for every real number $a>0$ we have

$$P(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Markov's Inequality

- Let X be a random variable that takes only nonnegative values. Then, for every real number $a>0$ we have

$$P(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

How large a value can X take?

Markov's Inequality: Proof

$$
\begin{aligned}
\mathbb{E}(X) &= \textstyle\sum_x xP(x) \\
&= \textstyle\sum_{x<a} xP(x) + \sum_{x\geq a} xP(x) \\
&\geq \textstyle\sum_{x<a} 0P(x) + \sum_{x\geq a} aP(x) \\
&= a\textstyle\sum_{x\geq a} P(x) \\
&= aP(x \geq a)
\end{aligned}
$$

Markov's Inequality: An Example

- A fair coin is tossed *n* times. Give an upper bound on the probability that at least *3n/4* of the tosses yield heads.

$$P(X \geq \frac{3n}{4}) \leq \frac{\mathbb{E}(X)}{3n/4} = \frac{n/2}{3n/4} = \frac{2}{3}$$

- For distributions encountered in practice, Markov's inequality gives a very loose bound.
  - Why?

Chebyshev's Inequality

- Let *X* be a random variable. For every real number *r>0,*

$$P(|X - \mathbb{E}(X)| \geq a) \leq \frac{V(X)}{a^2}$$

Chebyshev's Inequality

- Let X be a random variable. For every real number $r>0$,

$$P(|X - \mathbb{E}(X)| \geq a) \leq \frac{V(X)}{a^2}$$

How likely is it that RV X takes a value
that's at least distance a from its expected value?

Chebyshev's Inequality: Proof

Chebyshev's Inequality: Proof

Observe that

$$P(|X - \mathbb{E}(X)| \geq r) = P((X - \mathbb{E}(X))^2 \geq r^2)$$

Chebyshev's Inequality: Proof

Observe that

$$P(|X - \mathbb{E}(X)| \geq r) = P((X - \mathbb{E}(X))^2 \geq r^2)$$

Applying Markov's inequality, we get

$$P((X - \mathbb{E}(X))^2 \geq r^2) \leq \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{r^2} = \frac{V(X)}{r^2}$$

Markov vs Chebyshev

$$P(X \geq k\mu) \leq \frac{1}{k}$$

vs

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Chebyshev's Inequality: An Example

- Assume we have a distribution whose mean is 80 and standard deviation is 10. What is a lower bound on the percentage of values that fall between 60 and 100 (exclusively) in this distribution?

Chebyshev's Inequality: An Example

- Assume we have a distribution whose mean is 80 and standard deviation is 10. What is a lower bound on the percentage of values that fall between 60 and 100 (exclusively) in this distribution?

$$p(|X(s) - \mathbb{E}(X)| \geq r) \leq \frac{V}{r^2}$$

Chebyshev's Inequality: An Example

- Assume we have a distribution whose mean is 80 and standard deviation is 10. What is a lower bound on the percentage of values that fall between 60 and 100 (exclusively) in this distribution?

$$p(|X(s) - \mathbb{E}(X)| \geq r) \leq \frac{V}{r^2}$$

$$\mathbb{E}(X) = 80 \qquad V = 100 \qquad r = 20$$

Chebyshev's Inequality: An Example

- Assume we have a distribution whose mean is 80 and standard deviation is 10. What is a lower bound on the percentage of values that fall between 60 and 100 (exclusively) in this distribution?

$$p(|X(s) - \mathbb{E}(X)| \geq r) \leq \frac{V}{r^2}$$

$$\mathbb{E}(X) = 80 \qquad V = 100 \qquad r = 20$$

$$p(|X(s) - 80| \geq 20) \leq \frac{1}{4}$$

Chebyshev's Inequality: An Example

- Assume we have a distribution whose mean is 80 and standard deviation is 10. What is a lower bound on the percentage of values that fall between 60 and 100 (exclusively) in this distribution?

$$p(|X(s) - \mathbb{E}(X)| \geq r) \leq \frac{V}{r^2}$$

$$\mathbb{E}(X) = 80 \qquad V = 100 \qquad r = 20$$

$$p(|X(s) - 80| \geq 20) \leq \frac{1}{4}$$

$$\Rightarrow \text{ lower bound is } 75\%$$

Illustration: Estimating π Using the Monte Carlo Method

- Here's a simple algorithm for estimating π:
  - Throw darts at a square whose area is 1, inside which there's a circle whose radius is 1/2.
  - The probability that it lands inside the circle equals the ratio of the circle area to the square area ($\pi/4$). Therefore, calculate the proportion of times that the dart landed inside the circle and multiply it by 4.

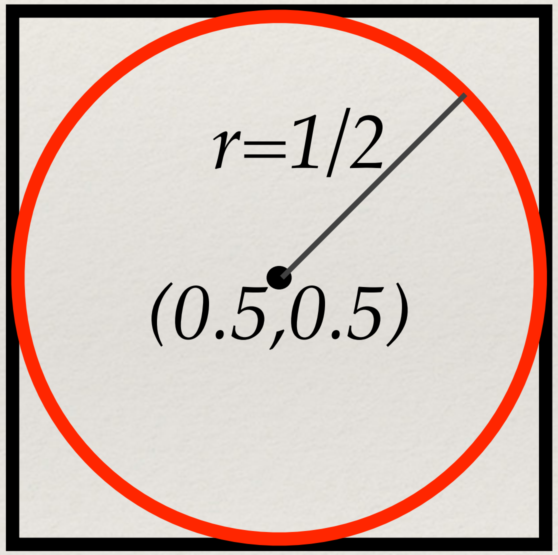


Illustration: Estimating π Using the Monte Carlo Method

Algorithm 1: MonteCarlo_πEstimation.

Input: $n \in \mathbb{N}$.
Output: Estimate $\hat{\pi}$ of π.

for $i = 1$ **to** n **do**
    $a \leftarrow random(0,1)$; `// random number in` $[0,1]$
    $b \leftarrow random(0,1)$; `// random number in` $[0,1]$
    $X_i \leftarrow 0$;
    if $\sqrt{(a-0.5)^2 + (b-0.5)^2} \leq 0.5$ **then**
        $X_i \leftarrow 1$; `// the dart landed inside/on the circle`

$\hat{\pi} \leftarrow 4 \cdot (\sum_{i=1}^{n} X_i)/n$;
return $\hat{\pi}$;

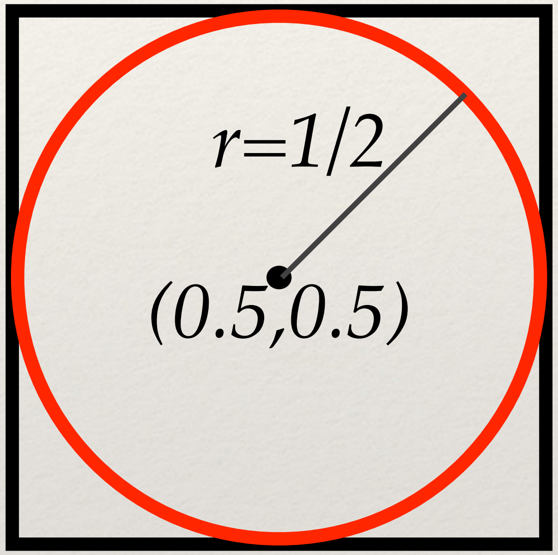


Illustration: Estimating π Using the Monte Carlo Method

- Let X_i be the random variable that denotes whether the i-th dart landed inside the circle (1 if it did, and 0 otherwise).
- Then, $\hat{\pi}(n) = 4\frac{\sum_{i=1}^{n} X_i}{n}$

Illustration: Estimating π Using the Monte Carlo Method

- Let X_i be the random variable that denotes whether the i-th dart landed inside the circle (1 if it did, and 0 otherwise).
- Then, $\hat{\pi}(n) = 4\frac{\sum_{i=1}^{n} X_i}{n}$

$$\mathbb{E}(X_i) = \frac{\pi}{4}1 + (1 - \frac{\pi}{4})0 = \frac{\pi}{4}$$

$$V(X_i) = \frac{\pi}{4}(1 - \frac{\pi}{4})$$

Illustration: Estimating π Using the Monte Carlo Method

- Let X_i be the random variable that denotes whether the i-th dart landed inside the circle (1 if it did, and 0 otherwise).
- Then, $\hat{\pi}(n) = 4\frac{\sum_{i=1}^{n} X_i}{n}$

$$\mathbb{E}(X_i) = \frac{\pi}{4}1 + (1 - \frac{\pi}{4})0 = \frac{\pi}{4}$$

$$V(X_i) = \frac{\pi}{4}(1 - \frac{\pi}{4})$$

$$\mathbb{E}(\hat{\pi}) = \mathbb{E}\left(\frac{4}{n}\sum_{i=1}^{n} X_i\right) = \frac{4}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \pi$$

Illustration: Estimating π Using the Monte Carlo Method

- Let X_i be the random variable that denotes whether the i-th dart landed inside the circle (1 if it did, and 0 otherwise).
- Then, $\hat{\pi}(n) = 4\frac{\sum_{i=1}^{n} X_i}{n}$

$$\mathbb{E}(X_i) = \frac{\pi}{4}1 + (1 - \frac{\pi}{4})0 = \frac{\pi}{4}$$

$$V(X_i) = \frac{\pi}{4}(1 - \frac{\pi}{4})$$

$$\mathbb{E}(\hat{\pi}) = \mathbb{E}\left(\frac{4}{n}\sum_{i=1}^{n} X_i\right) = \frac{4}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \pi$$

$$V(\hat{\pi}) = V(\frac{4}{n}\sum_{i=1}^{n} X_i) = \frac{16}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{\pi(4-\pi)}{n}$$

Illustration: Estimating π Using the Monte Carlo Method

- The question of interest is: How big should n be for us to get a good estimate?

Illustration: Estimating π Using the Monte Carlo Method

- In a probabilistic setting, the question can be asked as:
  - What should the value of n be so that the estimation error of π is within δ with probability at least ε?
  - (of course, we want δ to be very small and ε to be as close to 1 as possible. For example, δ=0.001 and ε=0.95)

Illustration: Estimating π Using the Monte Carlo Method

- In other words, we are interested in the value of n that yields

$$p(|\hat{\pi}(n) - \pi| < \delta) > \varepsilon$$

(equivalently, $p(|\hat{\pi}(n) - \pi| \geq \delta) \leq 1 - \varepsilon$)

Illustration: Estimating π Using the Monte Carlo Method

- For δ=0.001 and ε=0.95, we seek *n* such that

$$p(|\hat{\pi}(n) - \pi| \geq 0.001) \leq 0.05$$

Illustration: Estimating π Using the Monte Carlo Method

- For δ=0.001 and ε=0.95, we seek n such that

$$p(|\hat{\pi}(n) - \pi| \geq 0.001) \leq 0.05$$

Chebyshev's inequality: $\hat{\pi}(n)$ $\mathbb{E}(\hat{\pi})$ *r* *V*/*r*²

Illustration: Estimating π Using the Monte Carlo Method

- For δ=0.001 and ε=0.95, we seek n such that

$$p(|\hat{\pi}(n) - \pi| \geq 0.001) \leq 0.05$$

Chebyshev's inequality: $\hat{\pi}(n)$ $\mathbb{E}(\hat{\pi})$ *r* *V*/*r*²

$$V(\hat{\pi}(n)) = \frac{\pi(4-\pi)}{n}$$

Illustration: Estimating π Using the Monte Carlo Method

- For δ=0.001 and ε=0.95, we seek *n* such that

$$p(|\hat{\pi}(n) - \pi| \geq 0.001) \leq 0.05$$

Chebyshev's inequality: $\hat{\pi}(n)$ $\mathbb{E}(\hat{\pi})$ *r* *V*/*r*²

$$V(\hat{\pi}(n)) = \frac{\pi(4-\pi)}{n}$$

So, we would like *n* such that $\frac{\pi(4-\pi)}{n(0.001)^2} \leq 0.05$

Illustration: Estimating π Using the Monte Carlo Method

$$\frac{\pi(4-\pi)}{n(0.001)^2} \leq 0.05$$

Illustration: Estimating π Using the Monte Carlo Method

$$\frac{\pi(4-\pi)}{n(0.001)^2} \leq 0.05$$

$$\pi(4-\pi) \leq 4$$

Illustration: Estimating π Using the Monte Carlo Method

$$\frac{\pi(4-\pi)}{n(0.001)^2} \leq 0.05$$

$$\pi(4-\pi) \leq 4$$

$$\Rightarrow \frac{\pi(4-\pi)}{n(0.001)^2} \leq \frac{4}{n(0.001)^2} \leq 0.05$$

Illustration: Estimating π Using the Monte Carlo Method

$$\frac{\pi(4-\pi)}{n(0.001)^2} \leq 0.05$$

$$\pi(4-\pi) \leq 4$$

$$\Rightarrow \frac{\pi(4-\pi)}{n(0.001)^2} \leq \frac{4}{n(0.001)^2} \leq 0.05$$

$$\Rightarrow n \geq 80,000,000$$

A Corollary of Chebyshev's Inequality

- Let $X_1, X_2, \ldots, X_n$ be independent random variables with

$$\mathbb{E}(X_i) = \mu_i \qquad \text{and} \qquad V(X_i) = \sigma_i^2$$

Then, for any $a>0$:

$$P\left(\left|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right| \geq a\right) \leq \frac{\sum_{i=1}^{n} \sigma_i^2}{a^2}$$

The Law of Large Numbers

- Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed (i.i.d.) random variables, where the (unknown) expected value μ is the same for all variables (that is, $\mathbb{E}(X_i) = \mu$) and their variance is finite. Then, for any ε>0, we have

$$P\left(\left|\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right| \geq \varepsilon\right) \xrightarrow{n\to\infty} 0$$

Questions?