CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

Lecture 10:

- Geometric Distribution
  - Memoryless Property
- Negative Binomial

Review: Binomial Distribution

The Binomial Distribution occurs when you count the number of successes in N independent and identically distributed Bernoulli Trials (i.e., p is the same each time).

Formally, if Y ~ Bernoulli(p), and

X = "The number of successes in N trials of Y" = $\overbrace{Y + Y + \ldots + Y}^{\text{N times}}$

then we say that X is distributed according to the Binomial Distribution with parameters N and p, and write this as:

$$X \sim B(N, p)$$

where

$$R_X = \{\, 0, \ldots, N \,\}$$

$$f_X(k) = \binom{N}{k} p^k \, (1-p)^{n-k}$$

> Note: k successes and N-k failures:
>
> SSS...FFFF...
>
> has probability $p^k (1\text{-}p)^{N\text{-}k}$
>
> and there are C(N,k) such sequences.

and where

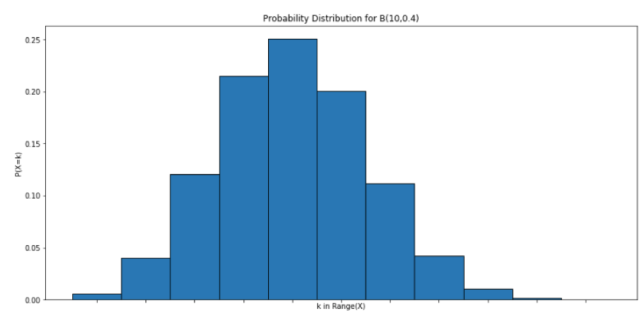
$$E(X) = N\,p \qquad Var(X) = N\,(1-p)\,p \qquad \sigma_X = \sqrt{N(1-p)p}$$

Problem: A restaurant serves 8 fish entrees, 12 beef, and 10 chicken. If customers select from these entrees randomly and independently, and the kitchen has plenty of each entree (so this is "with replacement"), what is the probability that exactly 4 of the next 10 customers order beef?

Solution:

N = 10 (number of customers who order)
p = 12/30 (probability that a random customer chooses beef)

X ~ B(10,12/30)



$$P(X = 4) = \binom{10}{4} \cdot \left(\tfrac{12}{30}\right)^4 \cdot \left(\tfrac{18}{30}\right)^6 = 0.2508$$

Geometric Distribution

Recall: The Geometric Distribution occurs when you count the number of independent and identically distributed Bernoulli trials until the first success.

Formally, if Y ~ Bernoulli(p), and

X = "The number of trials of Y until the first success"

then we say that X is distributed according to the Geometric Distribution with parameter p, and write this as:

$$X \sim G(p)$$

where

$$\begin{array}{rllllll} R_X = \{ & 1, & 2, & 3, & \ldots & k, & \ldots \} \\ S = \{ & S, & FS, & FFS, & \ldots & FFF..S, & \ldots \} \\ f_X = \{ & p, & (1-p)p, & (1-p)^2 p, & \ldots & (1-p)^{k-1} p, & \ldots \} \end{array}$$

For k, we have k-1 failures and 1 success (FFF... S), which has probability $(1\text{-}p)^{k\text{-}1}$ p .

Geometrical Distribution: G(p)

Motivation: This counts the number of Bernoulli trials until the first success occurs.

Definition and Example: X ~ G(p) if

$$
\begin{aligned}
Rng(X) &= \{1, 2, \ldots\} \\
f(k) &= (1-p)^{k-1} p
\end{aligned}
$$

$$
\begin{aligned}
E(X) &= \frac{1}{p} \\
Var(X) &= \frac{1-p}{p^2}
\end{aligned}
$$

$$
\begin{aligned}
P(X > k) &= (1-p)^k \\
P(X \leq k) &= 1.0 - (1-p)^k
\end{aligned}
$$

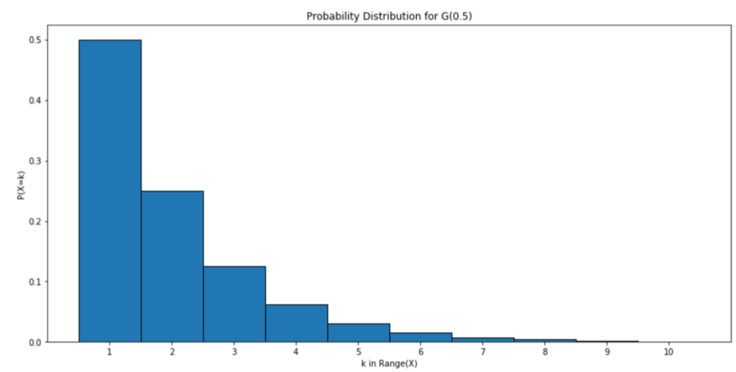
What is the probability when flipping a fair coin that it takes exactly 4 flips to get the first head?

$$P(X = 4) \ = \ (\tfrac{1}{2})^3(\tfrac{1}{2}) \ = \ \tfrac{1}{32}$$

Geometric Distribution: The Memoryless Property

The Geometric is one of two distributions that has the Memoryless Property, which we have already discussed informally before now as "the coin doesn't remember its past flips" but which is actually a much stronger statement about the entire distribution.

Intuitive Version: Suppose we are about to start flipping a coin for which the probability of heads is p. Then the probability distribution of X = "how many flips until the first head?" is G(p).



Now suppose that the first k flips are tails. Then the probability distribution of Y = "how many **more** flips until the first head?" is still G(p).

In other words, it doesn't matter when you start to count or what the past history is: the exact theoretical distribution is always the same.

Intuitive Proof: Suppose you come into the room while someone is flipping the coin. How do you know how many flips have occurred before you came in, and why would it matter? The distribution would be exactly the same irregardless of the past.

Geometric Distribution: The Memoryless Property

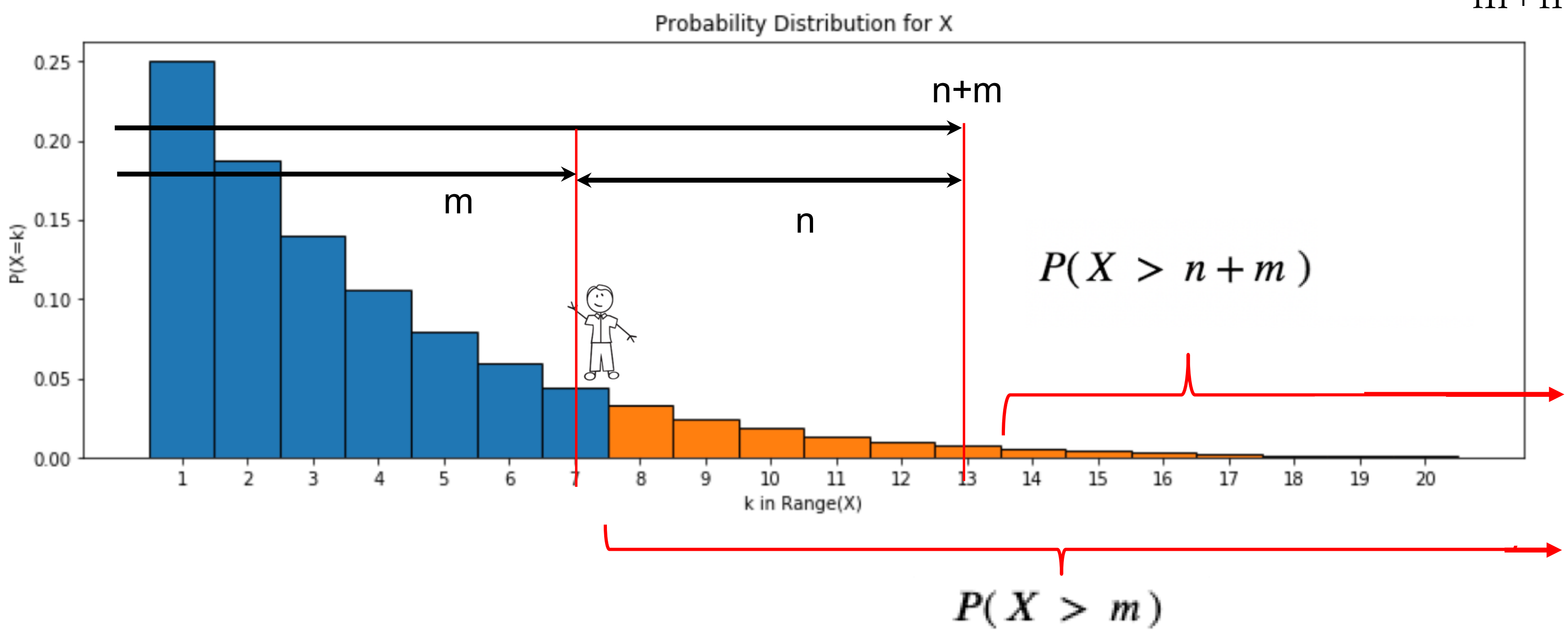
Theorem A random variable X is called memoryless if, for any n, m $\geq$ 0,

$$P(X > n + m \mid X > m) = P(X > n)$$

Example:

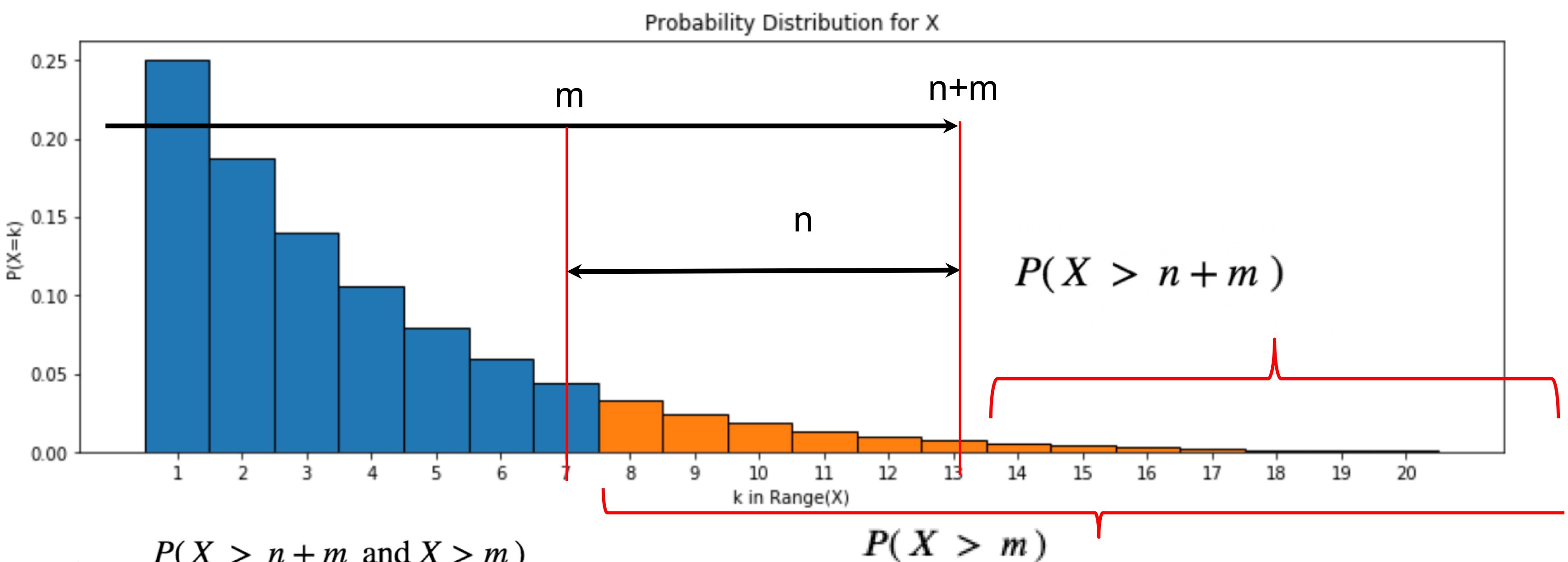
m = 7 n = 6
m+n = 13

Fact: For any probability p, X ~ G(p) has the memoryless property.



(In fact, the Geometric is the only discrete distribution with this property; a continuous version of the Geometric, called the Exponential, is the other one.)

Geometric Distribution: The Memoryless Property



Proof:

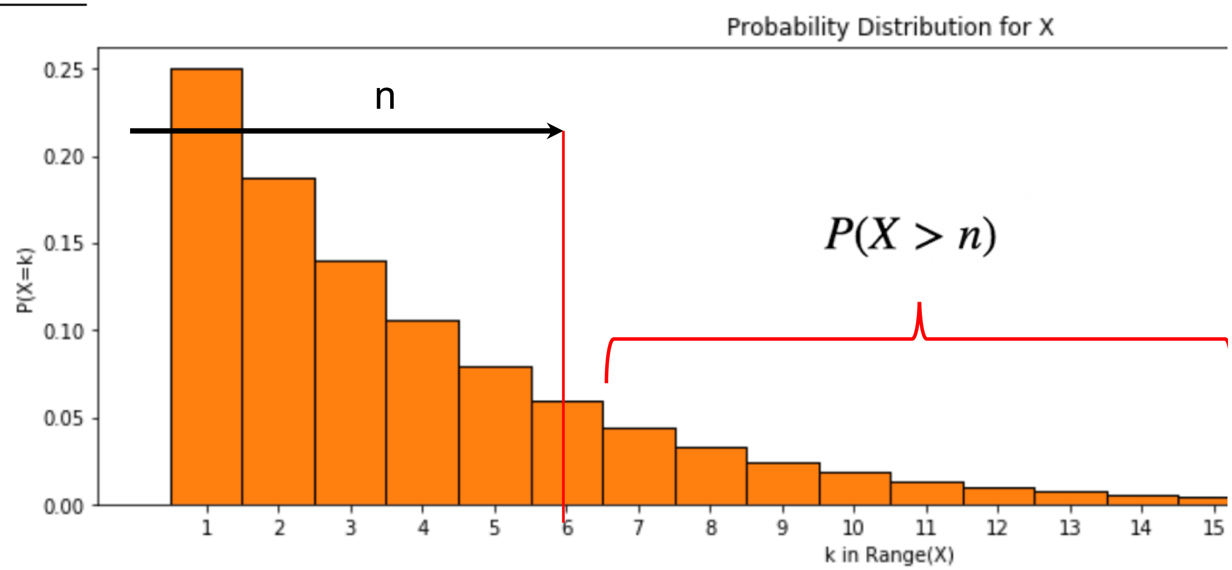
$$P(X > n+m \mid X > m) = \frac{P(X > n+m \text{ and } X > m)}{P(X > m)}$$

$$= \frac{P(X > n+m)}{P(X > m)}$$

$$= \frac{(1-p)^{(n+m)}}{(1-p)^{m}}$$

$$= (1-p)^{n}$$

$$= P(X > n)$$



Geometric Distribution: Expected Value and Variance

To derive the expected value, we can use the fact that X ~ G(p) has the memoryless property and break into two cases, depending on the result of the first Bernoulli trial. Let

X_S = "result of X when there is a success on the first trial"
X_F = "result of X when there is a failure on the first trial"

Clearly,

- $E(X_S) = 1$
- $E(X_F) = 1 + E($ X for the remaining trials $)$

  $= 1 + E(X)$

  By the memoryless property!

Thus we have:

$$
\begin{aligned}
\mu_X &= 1\,p + (1-p)\,(1+\mu_X) \\
&= p + 1 - p + \mu_X - p\,\mu_X \\
&= 1 + \mu_X - p\,\mu_X \\
0 &= 1 - p\,\mu_X \\
p\,\mu_X &= 1 \\
\mu_X &= 1/p
\end{aligned}
$$

Geometric Distribution

Example

Suppose you draw cards WITH replacement until you get an Ace. How many draws would you expect it to take?

Solution: This is $G(1/13)$. $E(X) = 13$

On average, how many independent games of poker are required until a particular player is dealt a **Royal Flush**?

Solution: This is $G(0.00000154)$. $E(X) = 1/ 0.00000154 = 649{,}350.6493$

Negative Binomial Distribution

A natural generalization of the Geometric is to ask how long before r >= 1 successes...

Formally, if Y ~ Bernoulli(p), and

X = "The number of trials of Y until the first r successes"

then we say that X is distributed according to the Negative Binomial Distribution with parameters r and p, and write this as:

$$X \sim NP(r,p)$$

where:

$$f_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

and:

$$E(X) = \frac{r}{p}$$

$$Var(X) = \frac{r(1-p)}{p^2}$$

This follows from the fact that before the r^{th} success, we had (r-1) successes among (k-1) Bernoulli trials, which is a Binomial:

$$\binom{k-1}{r-1} p^{r-1} (1-p)^{(k-1)-(r-1)} = \binom{k-1}{r-1} p^{r-1} (1-p)^{k-r}$$

Multiplying by p once more (for the last success) gives us our formula.

Negative Binomial Distribution

Example: Wayne and Liz play a series of backgammon games until one of them wins five games. Suppose that the games are independent and the probability that Liz wins any particular game is 0.58 (and Wayne's is therefore 0.42).

(a) What is the probability that the series ends after seven games?

(b) If the series does end in exactly seven games, what is the probability that Liz wins the series?

Solution: Let X be the number of games until Liz wins 5 games, and let Y be the number of games until Wayne wins 5 games. Thus,

$$X \sim NB(5, 0.58) \qquad Y \sim NB(5, 0.42)$$

(a)

$$P(X=7)+P(Y=7) = \binom{6}{4}(0.58)^5(0.42)^2 + \binom{6}{4}(0.42)^5(0.58)^2 = 0.17 + 0.066 = 0.24$$

(b)

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(X=7)}{P(X=7)+P(Y=7)} = \frac{0.17}{0.24} = 0.71$$