Geometric Distribution

(from last time)

- Def. A r.v. X has the geometric distribution with parameter p, $0 < p \le 1$, if $P(X=k) = (1-p)^{k-1}p$, k = 1,2,3,4…
  - Example: X could be the number of times you have to flip a coin before getting an H, if P(H) = p on any flip.
  - Note: the geometric distribution has infinitely many values, but is discrete.
- Theorem. If X is geometric with parameter p, then E(X) = 1/p, $V(X) = (1-p)/p^2$

Geometric Distribution

- let $f(x) = \Sigma_{n=0}^{\infty} x^n = (1-x)^{-1}$. Then:
- 1\. $f'(x) = \Sigma_{n=1}^{\infty} nx^{n-1} = (1-x)^{-2}$, and
- 2\. $f''(x) = \Sigma_{n=2}^{\infty} n(n-1)x^{n-2} = 2(1-x)^{-3}$.
- $E(X) = \Sigma_{n=1}^{\infty} nP(X=n) = \Sigma_{n=1}^{\infty} n(1-p)^{n-1}p$
  $= p(1-(1-p))^{-2} = 1/p$, using 1.
- $V(X) = \Sigma_{n=1}^{\infty} (n-p^{-1})^2 P(X=n)$
  $= \Sigma_{n=1}^{\infty} (n-p^{-1})^2 (1-p)^{n-1}p$
  $= \Sigma_{n=1}^{\infty} (n^2-2np^{-1}+p^{-2})(1-p)^{n-1}p$
  $= \Sigma_{n=1}^{\infty} (n(n-1) +n -2np^{-1}+p^{-2})(1-p)^{n-1}p$
  $= \Sigma_{n=1}^{\infty} (n(n-1) +n(1-2p^{-1})+p^{-2})(1-p)^{n-1}p$
  $= (1-p)p\Sigma_{n=2}^{\infty} n(n-1)(1-p)^{n-2} + (1-2p^{-1})p\Sigma_{n=1}^{\infty} n (1-p)^{n-1} + p^{-2} \Sigma_{n=1}^{\infty} (1-p)^{n-1}p$
  $=(1-p)p2p^{-3} + (p-2)p^{-2} + p^{-2}$, using 2, 1, & sum of all probs.
  $= p^{-2}(1-p)$

Chebyshev's Inequality

- Chebyshev's Inequality gives a bound on the probability that a random variable X, with sample space S, probability function p, takes on a value far from the mean, E(X).
- Theorem (p 491) (p 439 in 6th edition)
- $p(\{s : |X(s) – E(X)| \ge r\}) \le V(X)/r^2$

- Proof: Let $A = \{s : |X(s) – E(X)| \ge r\}$
- We need to show $p(A) \le V(X)/r^2$
- Now, $V(X) = \sum_{s\in S}(X(s) – E(X))^2 p(s)$
- $= \sum_{s\in A}(X(s) – E(X))^2 p(s) + \sum_{s\notin A}(X(s) – E(X))^2 p(s)$

$\ge \sum_{s\in A}(X(s) – E(X))^2 p(s)$

$\ge r^2 \sum_{s\in A} p(s) = r^2 p(A)$

since $|X(s) – E(X)| \ge r$ in A.

Now it's Time for…

Advanced Counting Techniques

Recurrence Relations

•A **recurrence relation** for the sequence $\{a_n\}$ is an equation that expresses a_n in terms of one or more of the previous terms of the sequence, namely, a_0, a_1, …, a_{n-1}, for all integers n with $n \geq n_0$, where n_0 is a nonnegative integer.

•A sequence is called a **solution** of a recurrence relation if its terms satisfy the recurrence relation.

Recurrence Relations

•In other words, a recurrence relation is like a recursively defined sequence, but **without specifying any initial values (initial conditions)**.

•Therefore, the same recurrence relation can have (and usually has) **multiple solutions**.

•If **both** the initial conditions and the recurrence relation are specified, then the sequence is **uniquely** determined.

Recurrence Relations

•**Example:**
Consider the recurrence relation
$a_n = 2a_{n-1} - a_{n-2}$ for n = 2, 3, 4, …

•Is the sequence $\{a_n\}$ with $a_n=3n$ a solution of this recurrence relation?

•For $n \geq 2$ we see that
$2a_{n-1} - a_{n-2} = 2(3(n - 1)) - 3(n - 2) = 3n = a_n$.

•Therefore, $\{a_n\}$ with $a_n=3n$ is a solution of the recurrence relation.

Recurrence Relations

- Is the sequence $\{a_n\}$ with $a_n=5$ a solution of the same recurrence relation?
- For $n \geq 2$ we see that $2a_{n-1} - a_{n-2} = 2 \cdot 5 - 5 = 5 = a_n$.
- Therefore, $\{a_n\}$ with $a_n=5$ is also a solution of the recurrence relation.

Modeling with Recurrence Relations

- **Example:**
- Someone deposits $10,000 in a savings account at a bank yielding 5% per year with interest compounded annually. How much money will be in the account after 30 years?

- **Solution:**
- Let P_n denote the amount in the account after n years.
- How can we determine P_n on the basis of P_{n-1}?

Modeling with Recurrence Relations

- We can derive the following **recurrence relation**:
- $P_n = P_{n-1} + 0.05P_{n-1} = 1.05P_{n-1}$.
- The initial condition is $P_0 = 10,000$.
- Then we have:
- $P_1 = 1.05P_0$
- $P_2 = 1.05P_1 = (1.05)^2P_0$
- $P_3 = 1.05P_2 = (1.05)^3P_0$
- …
- $P_n = 1.05P_{n-1} = (1.05)^nP_0$
- We now have a **formula** to calculate P_n for any natural number n and can avoid the iteration.

Modeling with Recurrence Relations

- Let us use this formula to find P_{30} under the
- initial condition $P_0 = 10,000$:

- $P_{30} = (1.05)^{30} \cdot 10,000 = 43,219.42$

•
After 30 years, the account contains $43,219.42.

Modeling with Recurrence Relations

•**Another example:**

•Let a_n denote the number of bit strings of length n that do not have two consecutive 0s ("valid strings"). Find a recurrence relation and give initial conditions for the sequence $\{a_n\}$.

•**Solution:**

•Idea: The number of valid strings equals the number of valid strings ending with a 0 plus the number of valid strings ending with a 1.

Modeling with Recurrence Relations

•Let us assume that $n \geq 3$, so that the string contains at least 3 bits.

•Let us further assume that we know the number a_{n-1} of valid strings of length (n – 1) and the number a_{n-2} of valid strings of length (n – 2).

•Then how many valid strings of length n are there, if the string ends with a 1?

•There are a_{n-1} such strings, namely the set of valid strings of length (n – 1) with a 1 appended to them.

•**Note:** Whenever we append a 1 to a valid string, that string remains valid.

Modeling with Recurrence Relations

•Now we need to know: How many valid strings of length n are there, if the string ends with a **0**?

•Valid strings of length n ending with a 0 **must have a 1 as their (n – 1)st bit** (otherwise they would end with 00 and would not be valid).

•And what is the number of valid strings of length (n – 1) that end with a 1?

•We already know that there are a_{n-1} strings of length n that end with a 1.

•Therefore, there are a_{n-2} strings of length (n – 1) that end with a 1.

Modeling with Recurrence Relations

•So there are a_{n-2} valid strings of length n that end with a 0 (all valid strings of length (n – 2) with 10 appended to them).

•As we said before, the number of valid strings is the number of valid strings ending with a 0 plus the number of valid strings ending with a 1.

•That gives us the following **recurrence relation**:

•$a_n = a_{n-1} + a_{n-2}$

Modeling with Recurrence Relations

- What are the **initial conditions**?
- $a_1 = 2$ (0 and 1)
- $a_2 = 3$ (01, 10, and 11)
- $a_3 = a_2 + a_1 = 3 + 2 = 5$
- $a_4 = a_3 + a_2 = 5 + 3 = 8$
- $a_5 = a_4 + a_3 = 8 + 5 = 13$
- …
- This sequence satisfies the same recurrence relation as the **Fibonacci sequence**.
- Since $a_1 = f_3$ and $a_2 = f_4$, we have $a_n = f_{n+2}$.

Solving Recurrence Relations

- In general, we would prefer to have an **explicit formula** to compute the value of a_n rather than conducting n iterations.
- For one class of recurrence relations, we can obtain such formulas in a systematic way.
- Those are the recurrence relations that express the terms of a sequence as **linear combinations** of previous terms.

Solving Recurrence Relations

- **Definition:** A linear homogeneous recurrence relation of degree k with constant coefficients is a recurrence relation of the form:
- $a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k}$,
- Where $c_1, c_2, \ldots, c_k$ are real numbers, and $c_k \neq 0$.
- A sequence satisfying such a recurrence relation is uniquely determined by the recurrence relation and the k initial conditions
- $a_0 = C_0, a_1 = C_1, a_2 = C_2, \ldots, a_{k-1} = C_{k-1}$.

Solving Recurrence Relations

- Why are they called linear?

If $\{x_n\}$ & $\{y_n\}$ are solutions of

$a_n = \Sigma_{j=1}^{k} c_j a_{n-j}$ then

$\{ux_n + vy_n\}$ is a solution, for u,v real.

- $ux_n = u\Sigma_{j=1}^{k} c_j x_{n-j} = \Sigma_{j=1}^{k} c_j ux_{n-j}$, and more generally,
- $ux_n+vy_n = u\Sigma_{j=1}^{k} c_j x_{n-j} + v\Sigma_{j=1}^{k} c_j y_{n-j}$
  $= \Sigma_{j=1}^{k} c_j(ux_{n-j} + vy_{n-i})$

Solving Recurrence Relations

•**Examples:**

•The recurrence relation $P_n = (1.05)P_{n-1}$
•is a linear homogeneous recurrence relation of **degree one**.

•The recurrence relation $f_n = f_{n-1} + f_{n-2}$
•is a linear homogeneous recurrence relation of **degree two**.

•The recurrence relation $a_n = a_{n-5}$
•is a linear homogeneous recurrence relation of **degree five**.

Solving Recurrence Relations

•Basically, when solving such recurrence relations, we try to find solutions of the form $\mathbf{a_n = r^n}$, where r is a constant.

•$a_n = r^n$ is a solution of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2} + \ldots + c_ka_{n-k}$ if and only if

•$r^n = c_1r^{n-1} + c_2r^{n-2} + \ldots + c_kr^{n-k}$.

•Divide this equation by r^{n-k} and subtract the right-hand side from the left:

•$r^k - c_1r^{k-1} - c_2r^{k-2} - \ldots - c_{k-1}r - c_k = 0$

•This is called the **characteristic equation** of the recurrence relation.

Solving Recurrence Relations

•The solutions of this equation are called the **characteristic roots** of the recurrence relation.

•Let us consider linear homogeneous recurrence relations of **degree two**.

•**Theorem:** Let c_1 and c_2 be real numbers. Suppose that $r^2 - c_1r - c_2 = 0$ has two distinct roots r_1 and r_2.

•Then the sequence $\{a_n\}$ is a solution of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$ if and only if $a_n = \alpha_1r_1^n + \alpha_2r_2^n$ for n = 0, 1, 2, …, where α_1 and α_2 are constants.

•See page 515 (6th Edition: pp. 414 and 415) for the proof.

Solving Recurrence Relations

•**Example:** What is the solution of the recurrence relation $a_n = a_{n-1} + 2a_{n-2}$ with $a_0 = 2$ and $a_1 = 7$?

•**Solution:** The characteristic equation of the recurrence relation is $r^2 - r - 2 = 0$.

•Its roots are r = 2 and r = -1.

•Hence, the sequence $\{a_n\}$ is a solution to the recurrence relation if and only if:

•$a_n = \alpha_1 2^n + \alpha_2(-1)^n$ for some constants α_1 and α_2.

Solving Recurrence Relations

- Given the equation $a_n = \alpha_1 2^n + \alpha_2(-1)^n$ and the initial conditions $a_0 = 2$ and $a_1 = 7$, it follows that
- $a_0 = 2 = \alpha_1 + \alpha_2$
- $a_1 = 7 = \alpha_1 \cdot 2 + \alpha_2 \cdot (-1)$
- Solving these two equations gives us $\alpha_1 = 3$ and $\alpha_2 = -1$.
- Therefore, the solution to the recurrence relation and initial conditions is the sequence $\{a_n\}$ with
- $a_n = 3 \cdot 2^n - (-1)^n$.

Solving Recurrence Relations

- **Another Example:** Give an explicit formula for the Fibonacci numbers.
- **Solution:** The Fibonacci numbers satisfy the recurrence relation $f_n = f_{n-1} + f_{n-2}$ with initial conditions $f_0 = 0$ and $f_1 = 1$.
- The characteristic equation is $r^2 - r - 1 = 0$.
- Its roots are

Solving Recurrence Relations

- Therefore, the Fibonacci numbers are given by

for some constants α_1 and α_2.
We can determine values for these constants so that the sequence meets the conditions $f_0 = 0$ and $f_1 = 1$:

Solving Recurrence Relations

- The unique solution to this system of two equations and two variables is

So finally we obtained an explicit formula for the Fibonacci numbers:

Solving Recurrence Relations

•But what happens if the characteristic equation has only one root?

•How can we then match our equation with the initial conditions a_0 and a_1 ?

•**Theorem:** Let c_1 and c_2 be real numbers with $c_2 \neq 0$. Suppose that $r^2 – c_1r – c_2 = 0$ has only one root r_0. A sequence $\{a_n\}$ is a solution of the recurrence relation $a_n = c_1a_{n-1} + c_2a_{n-2}$ if and only if $a_n = \alpha_1 r_0^n + \alpha_2 n r_0^n$, for n = 0, 1, 2, …, where α_1 and α_2 are constants. (Theorem 2, page 517)

13 Oct 2015 cs 320 29

Solving Recurrence Relations

•**Example:** What is the solution of the recurrence relation $a_n = 6a_{n-1} – 9a_{n-2}$ with $a_0 = 1$ and $a_1 = 6$?

•**Solution:** The only root of $r^2 – 6r + 9 = 0$ is $r_0 = 3$. Hence, the solution to the recurrence relation is

•$a_n = \alpha_1 3^n + \alpha_2 n3^n$ for some constants α_1 and α_2.

•To match the initial condition, we need

•$a_0 = 1 = \alpha_1$
$a_1 = 6 = \alpha_1 \cdot 3 + \alpha_2 \cdot 3$

•Solving these equations yields $\alpha_1 = 1$ and $\alpha_2 = 1$.

•Consequently, the overall solution is given by

•$a_n = 3^n + n3^n$.

13 Oct 2015 cs 320 30

Divide-and-Conquer Recurrences

•Some algorithms take a problem and **successively divide** it into one or more smaller problems until there is a **trivial solution** to them.

•For example, the **binary search** algorithm recursively divides the input into two halves and eliminates the irrelevant half until only one relevant element remained.

•This technique is called **"divide and conquer".**

•We can use **recurrence relations** to analyze the complexity of such algorithms.

13 Oct 2015 cs 320 31

Divide-and-Conquer Recurrences

•Suppose that an algorithm divides a problem (input) of size **n** into **a** subproblems, where each subproblem is of size **n/b**. Assume that **g(n)** operations are performed for such a division of a problem.

•Then, if **f(n)** represents the number of operations required to solve the problem, it follows that f satisfies the recurrence relation

•**$f(n) = af(n/b) + g(n)$**.

•This is called a **divide-and-conquer recurrence relation.**

13 Oct 2015 cs 320 32

Divide-and-Conquer Recurrences

- **Example:** The binary search algorithm reduces the search for an element in a search sequence of size **n** to the binary search for this element in a search sequence of size **n/2** (if n is even).
- **Two** comparisons are needed to perform this reduction.
- Hence, if **f(n)** is the number of comparisons required to search for an element in a search sequence of size n, then
- **f(n) = f(n/2) + 2** if n is even.