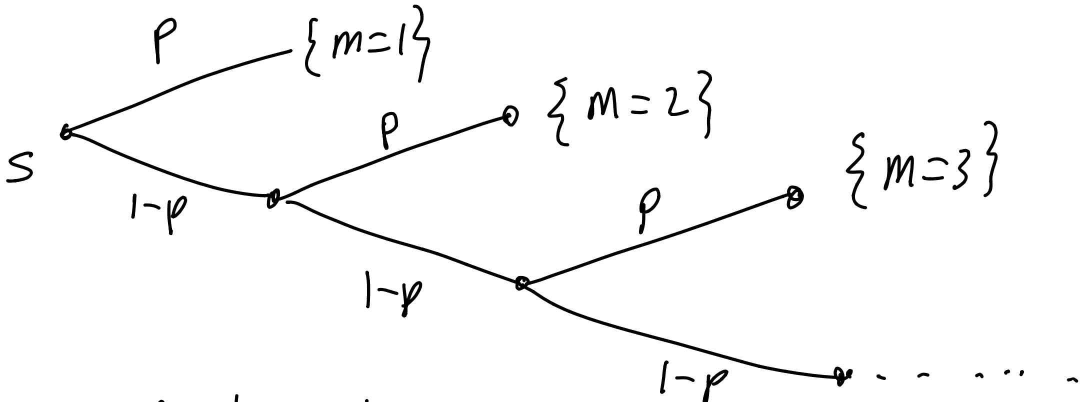
Geometric RV, parameter p

(Section 2.6.4 and 3.5.3)

The number of independent Bernouli trials until the first "success"

M is the RV, $M \in \{1, 2, 3, \ldots\}$



Dependent subexperiments
because the result of the previous subexperiment affects the current procedure.

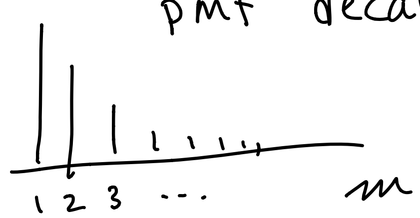
The outcome $M = m$ can only happen if the first $m-1$ trials were "failures" AND the m^{th} trial is a "success"

$$p_M(m) = \underbrace{(1-p)^{m-1}}_{m-1 \text{ failures}} \underbrace{p}_{\text{one success}}$$

Decays geometrically with m:

ratio $\dfrac{p_M(m+1)}{p_M(m)} = 1-p = q$

as p increases, pmf decays faster



Check : is this really a pmf?

want $$\sum_{m=1}^{\infty} p_M(m) = 1$$

$$\sum_{m=1}^{\infty} (1-p)^{m-1} p = p \sum_{m=1}^{\infty} (1-p)^{m-1}$$

change variables : $n = m-1$. n varies from 0 to ∞

$$= p \sum_{n=0}^{\infty} (1-p)^n = p\left(\frac{1}{1-(1-p)}\right) = \frac{p}{p} = 1$$

---

whats $P(M \le k)$?

$$= \sum_{j=1}^{k} p_M(j) = \sum_{j=1}^{k} p(1-p)^{j-1}$$

$$= p \sum_{i=0}^{k-1} (1-p)^i = p \, \frac{1-(1-p)^k}{1-(1-p)}$$

$$= 1-(1-p)^k$$

so $P(M > k) = (1-p)^k$ this is the same answer we'd get if we truncated the tree @ k.

Mean of geometric RV M

$$p_M(x) = p(1-p)^{x-1} \qquad x \in \{1, 2, \dots, \infty\}$$

$$E(M) = \sum_{k=1}^{\infty} k\,p(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1}$$

We know $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$ when $0<r<1$

differentiate both sides:

$$\sum_{k=0}^{\infty} k r^{k-1} = \frac{1}{(1-r)^2}$$

apply this to compute $E(m)$

$$E(m) = p\left[\frac{1}{1-(1-p)^2}\right] = \frac{p}{p^2} = \frac{1}{p}$$

$$Var(m) = E(m^2) - E(m)^2$$

using similar tools as above

$$E(m^2) = \frac{1+(1-p)}{p^2} = \frac{2-p}{p^2}$$

$$\text{so } Var(m) = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}$$

As p decreases, both the mean and variance increase

Be careful! There is another RV, also called the geometric

$m' = \#$ failures before 1^{st} success

$m' = m - 1$ (there's one more trial than failures)

$P(m'=k) = P(m=k+1) = (1-p)^k p \quad k = 0, 1, 2, \dots$

---

when applying the geometric, be very careful to identify the sample space — whether it starts at 0 or at 1, and carefully label your event $A = \{\text{"success"}\}$

Sometimes A is actually a failure!

Example: $A = \{\text{chip breaks when hit}\}$

Note: Since $m' = m - 1$

$$E(m') = E(m-1) = E(m) - 1$$

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

$$\text{Var}(m') = \text{Var}(m-1) = \text{Var}(m) = \frac{1-p}{p^2}$$

3.5 IMPORTANT DISCRETE RANDOM VARIABLES

Certain random variables arise in many diverse, unrelated applications. The pervasiveness of these random variables is due to the fact that they model fundamental mechanisms that underlie random behavior. In this section we present the most important of the discrete random variables and discuss how they arise and how they are interrelated. Table 3.1 summarizes the basic properties of the discrete random variables discussed in this section. By the end of this chapter, most of these properties presented in the table will have been introduced.

TABLE 3.1 Discrete random variables

Bernoulli Random Variable

$S_X = \{0, 1\}$

$p_0 = q = 1 - p \qquad p_1 = p \qquad 0 \le p \le 1$

$E[X] = p \quad \text{VAR}[X] = p(1 - p) \qquad G_X(z) = (q + pz)$

Remarks: The Bernoulli random variable is the value of the indicator function I_A for some event A; $X = 1$ if A occurs and 0 otherwise.

Binomial Random Variable

$S_X = \{0, 1, \ldots, n\}$

$$p_k = \binom{n}{k} p^k (1 - p)^{n-k} \qquad k = 0, 1, \ldots, n$$

$E[X] = np \quad \text{VAR}[X] = np(1 - p) \qquad G_X(z) = (q + pz)^n$

Remarks: X is the number of successes in n Bernoulli trials and hence the sum of n independent, identically distributed Bernoulli random variables.

Geometric Random Variable

First Version: $S_X = \{0, 1, 2, \ldots\}$

$p_k = p(1 - p)^k \qquad k = 0, 1, \ldots$

$$E[X] = \frac{1 - p}{p} \qquad \text{VAR}[X] = \frac{1 - p}{p^2} \qquad G_X(z) = \frac{p}{1 - qz}$$

Remarks: X is the number of failures before the first success in a sequence of independent Bernoulli trials. The geometric random variable is the only discrete random variable with the memoryless property.

Second Version: $S_{X'} = \{1, 2, \ldots\}$

$p_k = p(1 - p)^{k-1} \qquad k = 1, 2, \ldots$

$$E[X'] = \frac{1}{p} \qquad \text{VAR}[X'] = \frac{1 - p}{p^2} \qquad G_{X'}(z) = \frac{pz}{1 - qz}$$

Remarks: $X' = X + 1$ is the number of trials until the first success in a sequence of independent Bernoulli trials.

(*Continued*)

Closed-book Quiz 3

Fall 2016, TTh 3-4:15pm
(October 6, 2016)

Write your answers and your NAME clearly and legibly. Include your PUID. **THERE ARE THREE QUESTIONS. The maximum grade is a 2, but answer all 3 questions.**

Problem 1.
Mix and Match.
Each of the following world problems (items (1)-(4)) matches one and only one of the list of random variables (items (a)-(d)). For each world problem, identify the appropriate random variable.

1. Select and test integrated circuits until you find the first failure. Each test is independent with probability of failure 0.1. Which probability distribution best describes N, the random variable indicating the number of tests?
2. Select and test integrated circuits until you find the first failure. Each test is independent with probability 0.1 of finding a working IC. Which probability distribution best describes N, the random variable indicating the number of tests?
3. Select and test integrated circuits until you find the first failure. Each test is independent with probability 0.1 of finding a working IC. Which probability distribution best describes N, the random variable indicating the number of working ICs before finding a failure?
4. Select and test integrated circuits until you find the first failure. Each test is independent with probability of failure 0.1. Which probability distribution best describes N, the random variable indicating the number of working ICs before finding a failure?

(a) Geometric RV with $S = \{1, 2, 3, \ldots\}$ and $p = 0.1$

(b) Geometric RV with $S = \{0, 1, 2, \ldots\}$ and $p = 0.1$

(c) Geometric RV with $S = \{1, 2, 3, \ldots\}$ and $p = 0.9$

(d) Geometric RV with $S = \{0, 1, 2, \ldots\}$ and $p = 0.9$

Closed-book Quiz 3

Fall 2016, TTh 3-4:15pm
(October 6, 2016)

Write your answers and your NAME clearly and legibly. Include your PUID. **THERE ARE THREE QUESTIONS. The maximum grade is a 2, but answer all 3 questions.**

Problem 1.
Mix and Match.
Each of the following world problems (items (1)-(4)) matches one and only one of the list of random variables (items (a)-(d)). For each world problem, identify the appropriate random variable.

1. $p(\text{fail}) = 0.1$ Select and test integrated circuits until you find the first failure. Each test is independent with probability of failure 0.1. Which probability distribution best describes N, the random variable indicating the number of tests? A
2. $p(\text{good}) = 0.1$ Select and test integrated circuits until you find the first failure. Each test is independent with probability 0.1 of finding a working IC. Which probability distribution best describes N, the random variable indicating the number of tests? C
3. $p(\text{good}) = 0.1$ Select and test integrated circuits until you find the first failure. Each test is independent with probability 0.1 of finding a working IC. Which probability distribution best describes N, the random variable indicating the number of working ICs before finding a failure? D
4. $p(\text{fail}) = 0.1$ Select and test integrated circuits until you find the first failure. Each test is independent with probability of failure 0.1. Which probability distribution best describes N, the random variable indicating the number of working ICs before finding a failure? B

(a) Geometric RV with $S = \{1, 2, 3, \ldots\}$ and $p = 0.1$ 1

(b) Geometric RV with $S = \{0, 1, 2, \ldots\}$ and $p = 0.1$ 4

(c) Geometric RV with $S = \{1, 2, 3, \ldots\}$ and $p = 0.9$ 2

(d) Geometric RV with $S = \{0, 1, 2, \ldots\}$ and $p = 0.9$ 3

1	$A = \{\text{a failure}\}$	$P(A) = 0.1$	$S_N = \{1, \ldots\}$
2	"	$P(A) = 0.9$	$S_N = \{1, \ldots\}$
3	"	$P(A) = 0.9$	$S_N = \{0, 1, \ldots\}$
4	"	$P(A) = 0.1$	$S_N = \{0, 1, \ldots\}$

Example: Retransmissions (example 2.43 in Section 2.6.4)

Two computers communicate over an unreliable link.
The receiver can tell when an error occurs because of error detecting codes.
(Example: send 8 bits plus one parity bit)
If receiver detects an error, it asks the transmitter to send the message again.

Question: If the probability of a transmission error is $p = 0.1$, what is the probability a message is sent more than 2 times?

Answer: Independent Bernoulli trials until the message gets through.
$\Rightarrow$ failures until the 1st success, then stop
$\Rightarrow$ geometric.

$M = \{1, 2, \dots\}$ where M is # times message sent

$A = \{\text{transmission success}\}$

$P(A) = 1-p$ (Be careful! p was prob. failure)

So $P(M>2) = 1 - P(M=1) - P(M=2)$

$= 1 - p^0(1-p) - p(1-p) = p^2 = 0.01$