Geometric Distribution

Xinyu Liu, Qisheng Li, Leqi Zhao, Janice Zhang
November 28, 2016

Basics: Definition & Examples

A series of independent events with two outcomes, success (probability = p) and failure, define X to be the first success, X-1 is then the trails of failure before first success.

Examples:

Toss a coin repeatedly until the first head

Products produced by a machine has a 3% defective rate. What is the probability that the first defective occurs in the fifth item inspected?

P(X = 5) = P(1st 4 non-defective)P(5th defective) = (0.97^4) (0.03)

Pdf/cdf/MGF/E(x)/Variance/Example

Pdf: $\Pr(X = k) = (1-p)^{k-1}\,p$ (k = 1,2,...)

Cdf: $1-(1-p)^k$ (k = 1,2,...)

MGF = $\dfrac{pe^t}{1-(1-p)e^t}$ for $t < -\ln(1-p)$

E(x) = $\dfrac{1}{p}$

Variance = $\dfrac{1-p}{p^2}$

Relationship to Other Distributions

1. Exponential Distribution

$$If\ p = \frac{1}{n}\ and\ n \to \infty\ then\ we\ get\ an\ asymptotic$$
$$exponential\ distribution\ with\ parameter\ of\ \lambda = \frac{1}{n}$$

2. Negative Binomial Distribution

- Negative binomial distribution [X~NB (r, p)] describes the probability of x trials are made before r successes are obtained.
- Geometric distribution describes the probability of x trials a are made before one success.
- The geometric distribution *Y* is a special case of the negative binomial distribution, with *r* = 1

The Constant Rate Property

Suppose that T is a random variable taking values in ℕ+, which we interpret as the first time that some event of interest occurs.

The Rate Function of T is defined as:

$$h(n) = \mathbb{P}(T = n \mid T \geq n) = \frac{\mathbb{P}(T = n)}{\mathbb{P}(T \geq n)}, \quad n \in \mathbb{N}_+$$

R Code

Example

Application I: “While” Loops in Computer Science

1. Why it works?

- While Loop repeats the statements as long as some condition is TRUE (as long as the trial fails)
- While Loop breaks when this condition is FALSE (when we have the first success in trials)
- Statements in while loop are the Bernoulli trials. (e.g. random number generator can generate an outcome for each Bernoulli trials)

2. Evidence from Matlab:

https://www.geneseo.edu/~baldwin/csci120/spring2014/rec-while.html

Application II: St. Petersburg Paradox in Economics

1. Definition of St. Petersburg problem:

- Petersburg problem is a very famous game situation, in which a player bets on how many tosses of coins he/she will need before it first turns up heads.
- The player pays a fixed amount ($2) initially, and then receives 2^n dollars if the coin comes up heads on the nth toss.
- The game stops as soon as the player wins the game.

2. Why this works?

- We have n-1 tails (failure), and 1 head (success)
- P(tail) = ½ , P(head) = ½
- So, probability on the nth toss is P(fail for n-1 times U success for 1 time)
- Because of independence of trials, P(fail for n-1 times) * P(success for 1 time)
- P(nth toss) = $((\frac{1}{2})^{(n-1)}) * (1/2)$
- So, this is just geometric distribution with p = ½ .

Application III: Markov Chain

1. Definition

Markov Property (memoryless): if the probability of events happening in the future is independent of what went before, then the random variable is said to have the Markov property.

Markov Chain: a stochastic process that satisfies the Markov Property

It is applied in multiple fields: Computer Science, Physics, Chemistry, Statistics, Economics and Finance, Music

Application III: Markov Chain

2. Why this works?

Geometric distribution has Markov property

Example 1:

p = P(success)

q = 1 - p = P(fail)

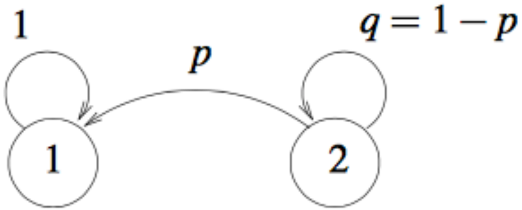


Figure 6: A Markov chain representing the geometric distribution.

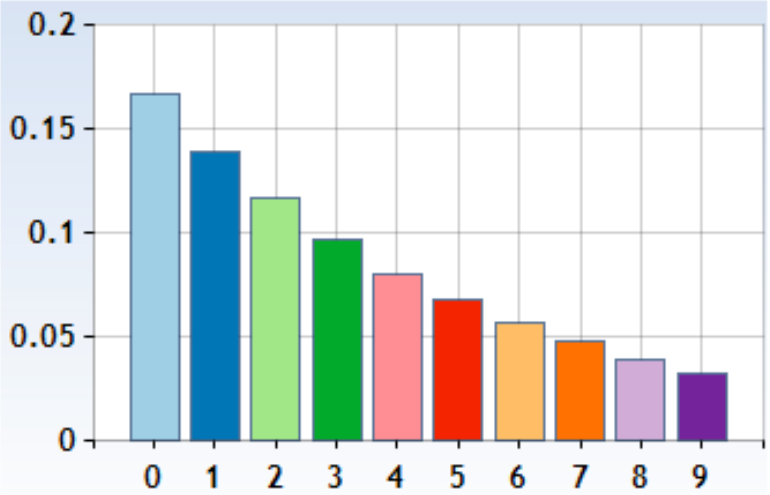
Example 2: Products are inspected until first defective is found. X is a geometric random variable with parameter p. The first 10 trials have been found to be free of defectives. What is the probability that the first defective will occur in the 15th trial?

Example I: Question

A die is rolled until a 1 occurs. What is the probability that we have to roll the die once, twice, and 3 times? Can you guess about the resulting geometric distribution?

Example 1: Solution

The probability of success of a single trial is $\frac{1}{6}$, so the above formula can be used directly:



The geometric distribution with $p = \frac{1}{6}$

$$\Pr(X = 1) = \frac{5^0}{6}\frac{1}{6} \approx .166$$

$$\Pr(X = 2) = \frac{5^1}{6}\frac{1}{6} \approx .139$$

$$\Pr(X = 3) = \frac{5^2}{6}\frac{1}{6} \approx .116$$

$$\Pr(X = 4) = \frac{5^3}{6}\frac{1}{6} \approx .096$$

$$\vdots$$

Reference

https://www.geneseo.edu/~baldwin/csci120/spring2014/rec-while.html

http://mathworld.wolfram.com/SaintPetersburgParadox.html

http://www-eio.upc.es/teaching/pe/www.computing.dcu.ie/~jhorgan/chapter10slides.pdf

https://brilliant.org/wiki/geometric-distribution/

http://www.math.uah.edu/stat/bernoulli/Geometric.html

http://www.tcs.hut.fi/Studies/T-79.250/tekstit/lecnotes_02.pdf

Q&A

?