Geometric Distributions

Dr. Bashar Al-Shboul

The University of Jordan

- The geometric distribution is the distribution produced by the random variable X defined to count the number of trials needed to obtain the first success.
  - - For example: Flipping a coin until you get a head or rolling a die until you get a 5
- A random variable X is geometric if the following conditions are met:
  - Each observation falls into one of just two categories, “success" or “failure."
  - The probability of success is the same for each observation.
  - The observations are all independent.
  - The variable of interest is the number of trials required to obtain the first success.

- Notice that this is different from the binomial distribution in that the number of trials is unknown.
- With geometric distributions we are trying to determine how many trials are needed in order to obtain a success.
- The probability that the first success occurs on the n^{th} trial is

$$P(X = n) = (1 - p)^{n-1}\, p$$

where *p* is the probability of success.

- The probability that it takes *more* than n trials to see the first success is

$$P(X> n) = (1- p)^n$$

- The mean, or expected number of trials to get a success in a geometric distribution is found by

$$E[X] = 1 / p$$

- And the variance is:

$$Var[X] = (1-p) / p^2$$

In MATLAB: geopdf & geocdf

geopdf

- P(X = k) = geopdf (k-1, p)

geocdf

- P(X <= k) = geocdf(k-1, p)
- P(X > k) = 1-geocdf (k-1, p)

- A quarter back completes 44% of his passes. We want to observe this quarterback during one game to see how many pass attempts he makes before completing one pass. **(p = 0.44)**
- a. What is the probability that the quarterback throws 3 incomplete passes before he has a completion ? **$P(X=4) = (1-p)^3p = 0.0773$**
- How many passes can the quarterback expect to throw before he completes a pass ? **$E[X] = (1/p) = 2.27$**
- Determine the probability that it takes more than 5 attempts before he completes a pass ? **$P(X > 5) = (1-p)^5 = 0.0551$ OR 1 – geocdf(4,0.44)**
- What is the probability that he attempts more than 7 passes before he completes one? **$P(X > 7) = (1-p)^7 = 0.0173$ OR 1 – geocdf(6,0.44)**

- Newsweek in 1989 reported that 60% of young children have blood lead levels that could impair their neurological development. Assuming a random sample from the population of all school children at risk, find:
  - The probability that at least 5 children out of 10 in a sample taken from a school may have a blood lead level that may impair development.
    - **Binomial Distribution / Success Fail situation with number of experiments known**
    - **$P(x >= 5) = 1 – P(X<=4) = 1 – binopdf(4, 10, 0.6) = 0.8338$**
  - The probability you will need to test 10 children before finding a child with a blood lead level that may impair development.
    - **Geometric / Success Fail, Number of experiments unknown, looking for first success**
    - **$P(X=11) = (1-0.6)^{10}(0.6)$ = approx. 6.29×10^{-5}**
  - The probability you will need to test no more than 10 children before finding a child with a blood lead level that may impair development.
    - **Geometric**
    - **$P(X<=10) = 1 – P(X>10) = 1 – (1-0.6)^{10} = 0.9999$ OR geocdf(9, 0.6)**