Lecture 4

Discrete Probability Distribution (1)

Statistics for
Civil & Environmental Engineers

Bernoulli Distribution(1)

- **Description**
  - **there are only two possible outcomes, called a *success* or *failure***
  - **the probability of occurrence of a success (or a failure) is *constant***
  - **the probability of the event occurring is *independent* of the time and *independent* of the past history of occurrences**

Bernoulli Distribution(2)

- **Bernoulli pmf**
  - $p_X(x) = \Pr[X = x; p] = p^x (1-p)^{1-x}$ for $x = 0,1,\ 0 \le p \le 1$
    $= 0$ otherwise
- **Population parameters**
  - **Mean=**
  - **Variance=**

Binomial Distribution(1)

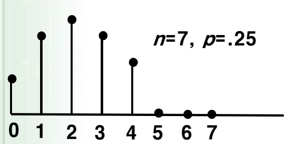
- **Description**
  - ***a series of Bernoulli trials*** **is made**
  - **the trials are conducted under** ***the same conditions***
  - **the order in which the events in the trials occur is immaterial**
- **Binomial pmf**
  - $p_X(x) = \Pr[X = x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0,1,..,n\ \ 0 \le p \le 1$
    $= 0$ otherwise
    where $n =$ number of trial, $x =$ number of successes
    $p =$ probability of success in any trial

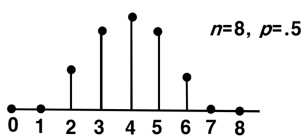
Binomial Distribution(2)

- Population parameters
  - Mean=
  - Variance=
  - Skewness coef.=
- Parameter estimates
  - $\hat{p} =$

Binomial Distribution(3)

- Comments
  - events must be independent
  - events must have only 2 possible outcomes (e.g., success or failure)
  - probability must be STABLE
  - for n=1, binomial reduces to Bernoulli distribution
- Distribution shapes





Binomial Distribution Example

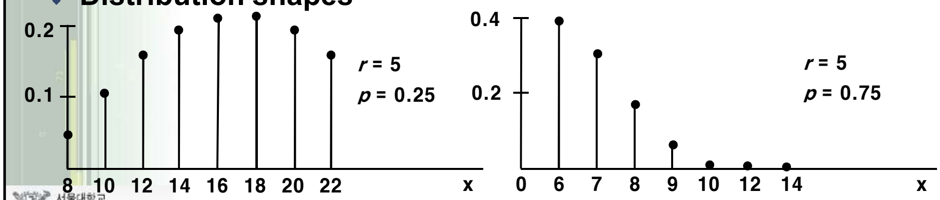
Example 4.2. Flooding of a road. Suppose a road is flooded with probability $p = .1$ during a year and not more than one flood occurs during a year. What is the probability that it will be flooded at least once during a five-year period?

Negative Binomial Distribution(1)

- **Description**
  - **a series of Bernoulli trials that are continued until exactly *r* success occur when X trials are required**
- **Negative Binomial pmf**
  - $$p_X(X = x, r, p) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad \text{for } x = r,\ r+1, \ldots$$
    $$= 0 \qquad \text{otherwise}$$
    where $x =$ number of trials, $r =$ number of successes
    $p =$ probability of success in any trial

Negative Binomial Distribution(2)

- **Population parameters**
  - **Mean=** $E[X] = r/p$
  - **Variance=** $Var[X] = r(1-p)/p^2$
- **Comments**
  - **for *r* =1, the negative binomial distribution reduces to the geometric distribution**
- **Distribution shapes**



Negative Binomial Distribution Example

Example 4.12. Delivery of treatment plants for a water supply system. A company has bid to supply standardized treatment plants for a rural water supply system in a region, having quoted a low price for the job. However the supervising engineer has estimated from previous experience that 10 percent of plants delivered are defective in some way. If five items are required, determine the minimum number of plants to be ordered to be 95 percent sure that a sufficient number of nondefective plants are delivered. It is assumed that the delivery of a plant is an independent trial and any fault that may occur in one plant is not related to possible faults in other plants. The probability of a success $p = 1 - .1 = .9$.

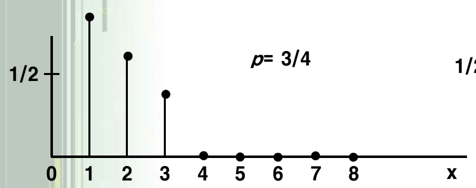
From Eq. (4.1.11) the cdf for $X = 7$ is given by

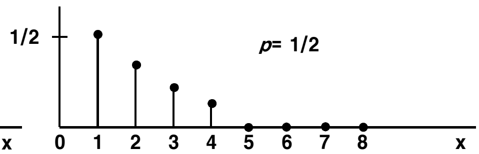
Geometric Distribution

- **Description**
  - **the probability that the first occurrence is at the *X*-th time**
- **Geometric pmf**
  - $$p_X(X=x;p)=p(1-p)^{x-1} \quad \text{for } x=1,2,3,\ldots,n \qquad 0\le p\le 1$$
    $$=0 \qquad \text{otherwise}$$
- **Population parameters**
  - **Mean=**
  - **Variance=**

Geometric Distribution

- **Comments**
  - **the probability distribution of the length of time between occurrences**
- **Distribution shapes**





Return Period

- **Definition**
  - **The mean time interval between exceedances of a specified value**
  - **The reciprocal of the probability of exceedance**
- **Equation**
  - $E(\tilde{T}) = \dfrac{1}{P_Y(Y > y)}$

Return Period Example

Example 4.7. In order to be 90 percent sure that a design storm is not exceeded in a 10-year period, what should be the return period of the design storm?

Hypergeometric Distribution(1)

- **Description**
  - **drawing a random sample of size *n* without replacement**
  - **from a finite population of size *m* with the elements of the population divided into two groups with *mp* elements belonging to one group**
- **Negative Binomial pmf**
  - $p_X(X=x)=\binom{mp}{x}\binom{m(1-p)}{n-x}\Big/\binom{m}{n}$ for $x=1,2,\ldots,\min(mp,n)$

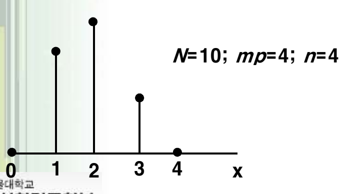
    where $x=$ number of successes in the sample

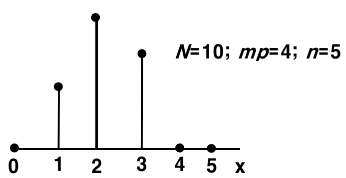
    $n=$ size of sample, $m=$ size of (finite) population

    $mp=$ number of successes in the population

Hypergeometric Distribution(2)

- **Population parameters**
  - **Mean=**
  - **Variance=**
- **Comments**
  - **if *n* is "small" compared to *m*, the binomial distribution is reasonable approximation for the hypergeometric distribution**
- **Distribution shapes**





Hypergeometric Distribution Example

Example 4.18. Components for water pumps. Standard pump components ordered for a water supply system all have the same specifications but p percent are found to be defective. A consignment of 100 items has been received. For this consignment to be accepted, no more than one item in a lot of 10 items selected at random can be defective. Compare the probabilities by the binomial and hypergeometric distributions, assuming $p = .02$.