**Business Statistics –– Discrete Probability Distributions**

|  |
| --- |
| Discrete Distributions, Empirical, Binomial, Poisson |

|  |  |  |
| --- | --- | --- |
|  | **Empirical Distribution** |  |

Suppose a questionnaire was given to 5 people containing the following question.

|  |
| --- |
| 1. How many miles did you travel to get here today? \_\_\_\_\_ |

Let an **Event** be one subject selected at random and observe the survey results of the subject.

Consider the likelihood or **probability** of an event.

If we select one of the subjects at random, consider probability statements about the miles the subject traveled?

Let the random variable, X=Miles

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Subject | Miles |  | X | Frequency | Probability |  |  |  | | | | | | | | | |
| 1 | 1 |  | 1 | 1 | 0.2 |  |  |  | | | | | | | | | |
| 2 | 3 |  | 3 | 2 | 0.4 |  |  |  | | | | | | | | | |
| 3 | 3 |  | 5 | 1 | 0.2 |  |  |  | | | | | | | | | |
| 4 | 5 |  | 8 | 1 | 0.2 |  |  |  | | | | | | | | | |
| 5 | 8 |  | Sum | 5 | 1.0 |  |  |  | | | | | | | | | |
|  |  |  |  |  |  |  |  |  | | | | | | | | | |
| Subject | Miles |  | X | Frequency | Probability |  | Probability | Probability Distribution | | | | | | | | | |
| 1 | 1 |  | 1 | 1 | 0.2 |  | 0.5 |  |  |  |  |  |  |  |  |  |  |
| 2 | 3 |  | 3 | 2 | 0.4 |  | 0.4 |  |  |  |  |  |  |  |  |  |  |
| 3 | 3 |  | 5 | 1 | 0.2 |  | 0.3 |  |  |  |  |  |  |  |  |  |  |
| 4 | 5 |  | 8 | 1 | 0.2 |  | 0.2 |  |  |  |  |  |  |  |  |  |  |
| 5 | 8 |  | Sum | 5 | 1.0 |  | 0.1 |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  | X |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Subject | Miles |  | X | Frequency | Probability |  | Frequency | Frequency Distribution | | | | | | | | | |
| 1 | 1 |  | 1 | 1 | 0.2 |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | 3 |  | 3 | 2 | 0.4 |  | 2 |  |  |  |  |  |  |  |  |  |  |
| 3 | 3 |  | 5 | 1 | 0.2 |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 5 |  | 8 | 1 | 0.2 |  | 1 |  |  |  |  |  |  |  |  |  |  |
| 5 | 8 |  | Sum | 5 | 1.0 |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  | X |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |  |

|  |  |  |
| --- | --- | --- |
| Probability Statements |  | Terminology: |
| P[X>4]=0.4 |  | Probability distribution or Probability Density Function (pdf) |
| P[X=5]=0.2 |  | Probability measure of a random variable, P[X>6]=0.2 |
| P[X>=3]=0.8 |  |  |
| P[1<X<4]=0.4 |  |  |
| P[1<=X<=4]=0.6 |  |  |

|  |  |  |
| --- | --- | --- |
|  | **Binomial Distribution** |  |

Binomial. X~B(n,p).

Definition.

Given a trial with two outcomes, success or failure.

Given a probability of success of an outcome of the trial is constant, p.

Given the outcome of the trials are independent.

Given a large number of such trials, n.

Given the number of successes in n trials, X.

Then, the probability of X is Binomial.

Examples.

1. If five people are selected at random from a population of half seniors, what is the probability that less than 2 people are seniors? n=5, p=0.5, X~B(n=5,p=0.5), P[X<2].

2. If only 10 employees out of 50 own compact cars, then what is the probability that at least one employee owns a compact car out of a random sample of three employees? n=3, p=10/50=0.2, X~B(n=3,p=0.2), P[X>=1].

3. Suppose a major at a university contains 20 sophomores, 15 juniors, and 10 seniors. What is the probability that a focus group of 12 students selected at random from the major contains no seniors?

n=12, p=(20+15)/(20+15+10)=7/9, X~B(n=12,p=7/9), P[X=0].

Calculations using Excel.

Using Excel, type “=BINOM.DIST(X,n,p,#)”, where “#=1 for density, #=2 for cumulative.”

Let X~B(n=5,p=0.5).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Find P[X<2]. Using Excel, type { =BINOM.DIST(1,5,0.5,1)=0.1875 }  Find P[X<=2]. Using Excel, type { =BINOM.DIST(2,5,0.5,1)=0.5 }  Find P[X=2]. Using Excel, type { =BINOM.DIST(2,5,0.5,0)=0.3125 }  Find P[X>2]. Using Excel, type { =1 – BINOM.DIST(2,5,0.5,1)=1 – 0.5 = 0.5 } | | | | |
|  |  |  |  |  |
| . . .    . . . | | | | |

|  |  |  |
| --- | --- | --- |
|  | **Binomial Examples** |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X~B(3,0.5) |  | X~B(3,0.2) |  | X~B(3,0.05) |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | pdf | PDF |  | pdf | PDF |  | pdf | PDF |
| X | f(X) | F(X) |  | f(X) | F(X) |  | f(X) | F(X) |
| 0 | .125 | .125 |  | .512 | .512 |  | .857 | .857 |
| 1 | .375 | .5 |  | .384 | .896 |  | .135 | .992 |
| 2 | .375 | .875 |  | .096 | .992 |  | .007 | .999 |
| 3 | .125 | 1 |  | .008 | 1 |  | .001 | 1 |
| Sum | 1 |  |  | 1 |  |  | 1 |  |

For each problem X~B(n,p), complete the sentence defining the terms:

X = number of \_\_\_\_\_\_\_\_\_ in n=\_\_\_\_\_ trials with p=\_\_\_\_\_\_\_\_\_\_

Successes size P[success]

1. If half of all shoppers at a mall are seniors, then from a random sample of three people, what is the probability that

|  |  |  |  |
| --- | --- | --- | --- |
| a. exactly one person is senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.375 | X~B(n=3, p=\_\_\_\_\_\_) |
| b. at least one person is senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.875 | X~B(n=3, p=\_\_\_\_\_\_) |
| c. at most two people are seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.875 | X~B(n=3, p=\_\_\_\_\_\_) |
| d. less than two people are seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.500 | X~B(n=3, p=\_\_\_\_\_\_) |
| e. greater than one person is not a senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.500 | X~B(n=3, p=\_\_\_\_\_\_) |
| f. no more than two people are not seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.875 | X~B(n=3, p=\_\_\_\_\_\_) |
| g. no less than one person is not a senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.875 | X~B(n=3, p=\_\_\_\_\_\_) |
| h. exactly two people are not seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.500 | X~B(n=3, p=\_\_\_\_\_\_) |

2. If 20% of all shoppers at a mall are senior, then from a random sample of three people, what is the probability that

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| a. exactly one person is senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.384 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| b. at least one person is senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.488 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| c. at most two people are seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.992 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| d. less than two people are seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.896 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| e. greater than one person is not a senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.896 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| f. no more than two people are not seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.488 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| g. no less than one person is not a senior? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.992 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| h. exactly two people are not seniors? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.384 | X~B(n=3, p=\_\_\_\_\_\_) |  |

3. If 5% of items are defective coming off an assembly line, then from a random sample of three items, what is the probability that

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| a. exactly one item is defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.135 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| b. at least one item is defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.143 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| c. at most two items are defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.999 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| d. less than two items are defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.992 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| e. greater than one item is not defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.992 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| f. no more than two items are not defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.143 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| g. no less than one item is not defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.999 | X~B(n=3, p=\_\_\_\_\_\_) |  |
| h. exactly two items are not defective? | P[X \_\_\_\_ \_\_\_\_\_\_\_]= | 0.135 | X~B(n=3, p=\_\_\_\_\_\_) |  |

Suppose a random sample of three items was repeated 3 times. What is the probability of no defectives found? [ (1-(1-0.05)3 )3 =0.003 ]

|  |  |  |
| --- | --- | --- |
|  | **Poisson Distribution** |  |

Poisson. X~Pssn().

Definition.

Let X represent the number of random, independent occurrences in an interval of consideration. The probability of X is Poisson where  is the average number of occurrences.

Examples.

1. If the average number of auto accidents in downtown on a particular holiday is 1.2/hour, then what is the probability that there are less than 2 accidents in any given hour? =1.2, X~Pssn(=1.2), P[X<2].

2. If the average number of clerical errors for a tax accounting firm is 0.7 for every tax return, then what is the probability that there are more than 2 errors out of 5 returns? =0.7\*5=3.5, X~Pssn(=3.5), P[X>2].

Calculations using Excel.

Using Excel, type “=Poisson.dist(X,,#)”, where “#=1 for density, #=2 for cumulative.”

Let X~Pssn(=0.6)

|  |
| --- |
| Find P[X<2]. Using Excel, type { =POISSON.DIST(1,0.6,1)=0.8781 }  Find P[X<=2]. Using Excel, type { =POISSON.DIST(2,0.6,1)=0.9769 }  Find P[X=2]. Using Excel, type { =POISSON.DIST(2,0.6,0)=0.0988 }  Find P[X>2]. Using Excel, type { =1–POISSON.DIST(2,0.6,1)=1–0.9769 = 0.0231 } |
|  |
| . . .    . . . |

|  |  |  |
| --- | --- | --- |
|  | **Poisson Examples** |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X~Poisson(0.1) |  | X~Poisson(0.2) |  | X~Poisson(0.3) |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | pdf | PDF |  | pdf | PDF |  | pdf | PDF |
| X | f(X) | F(X) |  | f(X) | F(X) |  | f(X) | F(X) |
| 0 | .9048 | .9048 |  | .8187 | .8187 |  | .7408 | .7408 |
| 1 | .0905 | .9953 |  | .1637 | .9824 |  | .2223 | .9631 |
| 2 | .0045 | .9998 |  | .0164 | .9988 |  | .0333 | .9964 |
| 3 | .0002 | 1 |  | .0011 | .9999 |  | .0033 | .9997 |
| 4 | .0000 | 1 |  | .0001 | 1 |  | .0003 | 1 |
| Sum | 1 |  |  | 1 |  |  | 1 |  |

1. Suppose call arriving at a switchboard follow a Poisson process. If the number of calls averages one call every 20 seconds, then what is the probability that

|  |  |  |
| --- | --- | --- |
| a. no calls in the first 2 seconds? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.9048 |
| b. at least one call in the first 4 seconds? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.1813 |
| c. no more than two calls in the first 6 seconds? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.9964 |
| d. no calls in the first minute? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.05 |

2. Suppose potholes in a county road follow a Poisson process. If the number of potholes averages one every 10 miles, then what is the probability that

|  |  |  |
| --- | --- | --- |
| a. no potholes in a randomly selected mile of road? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.9048 |
| b. at least one pothole in a random 3 miles of road? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.2592 |
| c. no more than two potholes in a random 2 miles stretch? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.9988 |
| d. at least one pothole within 500 feet either side  of a randomly selected junction? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.0188 |

3. Suppose the imperfections in the weave of a certain textile follow a Poisson process. If the average number of imperfections is one per five square yards, then what is the probability that

|  |  |  |
| --- | --- | --- |
| a. no imperfections in 1 yd2? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.8187 |
| b. at least one imperfections in ½ yd2? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.0952 |
| c. no more than 2 imperfections in 1.5 yd2? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.9964 |
| d. at least one imperfection in 5 square feet? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.1052 |

4. Suppose the occurrence of bacteria colonies of a certain type in a sample of polluted water follows a Poisson process. If the mean number of colonies is one in every ten cubic centimeters of water, then what is the probability that there will be

|  |  |  |
| --- | --- | --- |
| a. no bacteria colonies in 3 cc of water? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.7408 |
| b. at least one bacteria colony in 1 cc of water? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.0952 |
| c. no more than two bacteria colonies in 2 cc of water? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.9988 |
| d. at least one out of three 1-cc samples contain  at least one bacteria colony? | P[X \_\_\_\_ \_\_\_\_\_ | =\_\_\_\_\_]= | 0.2592 |

[ =0.1, P1=1-e-0.1=0.0952, P2=1-(1-P1)n=1-(1-(1- e-0.1))3=1- e-0.3 =0.2592 ]

|  |  |  |
| --- | --- | --- |
|  | **Binomial Distribution Summary** |  |
| . . .   |  | | --- | | X~B(n,p)  n=total number of trials  p=probability of success  f(X)= px (1–p)n–x  n!/(x!(n–x)!)  E[x] = np  V[x] = np(1–p)  For p<0.5, skewed right  For p=0.5, symmetric  For p>0.5, skewed left |   . . . | | |

|  |  |  |
| --- | --- | --- |
|  | **Poisson Distribution Summary** |  |
| . . .   |  | | --- | | X~Pssn()  =mean  f(X)= x e / x!  E[x] =   V[x] =   Always skewed right |   . . . | | |